

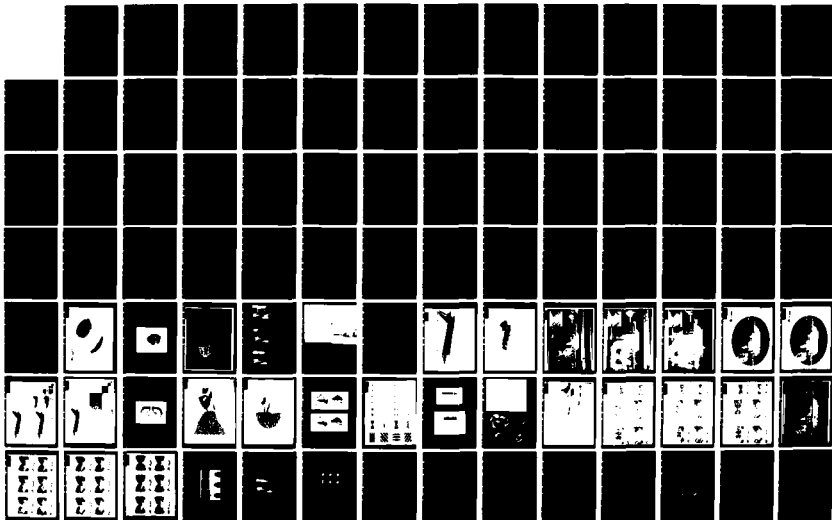
AD-A166 222

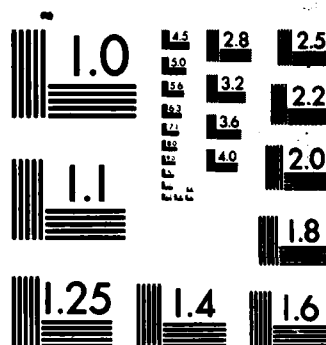
NOVEL ARCHITECTURES FOR IMAGE PROCESSING BASED ON  
COMPUTER SIMULATION AND (U) NEW YORK UNIV MEDICAL  
CENTER N Y E L SCHWARTZ 02 JAN 86 AFOSR-TR-86-0059  
F49620-83-C-0108 F/G 5/10

1/1

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A



**AFOSR-TR- 86-0059**

# NEW YORK UNIVERSITY MEDICAL CENTER

*A private university in the public service*

School of Medicine

550 FIRST AVENUE, NEW YORK, N.Y. 10016

CABLE ADDRESS: NYUMEDIC

Department of Psychiatry  
**BRAIN RESEARCH**  
(212) 340-

**AD-A166 222**

## FINAL REPORT

### TITLE OF PROJECT:

NOVEL ARCHITECTURES FOR IMAGE PROCESSING BASED ON COMPUTER  
SIMULATION AND PSYCHOPHYSICAL STUDIES OF HUMAN VISUAL CORTEX

### FUNDING AGENCY:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH IMAGE UNDERSTANDING  
PROGRAM.

CONTRACT F49620-83-C-0108

### PRINCIPAL INVESTIGATOR:

Dr. Eric L. Schwartz Phone: 212-340-6299  
S.S.#054-40-8944

Associate Professor of Psychiatry  
Dept. of Psychiatry  
Brain Research Laboratories  
New York University School of Medicine

### APPLICANT ORGANIZATION:

New York University School of Medicine  
550 First Avenue  
New York, N.Y. 10016

### ENTITY IDENTIFICATION NUMBER:

13-5562309A1

### PERIOD OF FUNDING:

4/15/83 - 4/15/85

**DTIC FILE COPY**

**DTIC  
ELECTE  
APR 1 1986  
D**

Approved for public release;  
distribution unlimited.

Author	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Editor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reviewer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Director	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Distribution/			
Availability Codes			
Avail and/or			
Special			
Dist			



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			5. MONITORING ORGANIZATION REPORT NUMBER <b>AFOSR-TR-86-0089</b>		
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research/NL		
6a. NAME OF PERFORMING ORGANIZATION New York University		6b. OFFICE SYMBOL (If applicable)		7b. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, DC 20332-6448	
6c. ADDRESS (City, State and ZIP Code) School of Medicine 550 First Avenue New York, NE 10016		8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		8b. OFFICE SYMBOL (If applicable) NL	
8c. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB DC 20332-6448		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER			
10. SOURCE OF FUNDING NOS					
PROGRAM ELEMENT NO. 61102F		PROJECT NO. 2313		TASK NO. A5	
WORK UNIT NO.					
11. TITLE (Include Security Classification) NOVEL ARCHITECTURES FOR IMAGE PROCESSING BASED ON COMPUTER SIMULATION AND PSYCHOPHYSICAL STUDY OF HUMAN VISUAL CORTIX					
12. PERSONAL AUTHOR(S) Professor Eric L. Schwartz					
13a. TYPE OF REPORT Final		13b. TIME COVERED FROM 83/4/15 TO 85/4/1		14. DATE OF REPORT (Yr., Mo., Day) 86 Jan 02	
15. PAGE COUNT 100					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB. GR.	VISION, PSYCHOPHYSICS, IMAGE PROCESSING		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)  This final report consists of two parts. The first part is a computer simulation of the functional architecture of visual cortex, and an examination of the possible significance that this architecture may have for understanding both human visual computation and machine vision. The second part of this report is a psychophysical investigation of human shape perception in terms of boundary descriptors of curvature. This progress report will deal with these two areas separately.					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL JOHN F TANGNEY		22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5021		22c. OFFICE SYMBOL NL	

MATTHEW J. KEMPER  
FINAL REPORT  
Chief, Technical Information Division

This final report consists of two parts. The first part is a computer simulation of the functional architecture of visual cortex, and an examination of the possible significance that this architecture may have for understanding both human visual computation and machine vision. The second part of this report is a psychophysical investigation of human shape perception in terms of boundary descriptors of curvature. This progress report will deal with these two areas separately.

I. Computer simulation of the functional architecture of primary visual cortex.

First, we will present a brief summary of the proposed research, followed by a statement of the progress to date.

A. Proposed research( Previous funding period)

1. Computer simulation of visual cortex

1.0 Hardware environment setup

Configure a 32-bit micro-processor ( M68000) system together with an image processing system which will allow image processing, solids modeling, and image data acquisition and display for gray-scale, motion, stereo, and color modalities.

1.1 Software environment setup

Construct software environment to provide basic image processing, solids modeling and simulation support.

1.2 Global topographic mapping of visual cortex

The space variant topographic mapping of primary visual cortex will be modeled using a global complex logarithmic mapping.

1.3 Space variant digital filtering

A space variant digital filter, following the curve of cortical magnification will be applied to the re-mapped point image.

1.4 Ocular dominance columns

A left and right eye view, interlaced according to the "ocular dominance column" pattern described by LeVay, Hubel and Weisel ( LeVay 1975 ) will be simulated , and used to

study stereo segmentation, according to previously published work ( Schwartz 1982 ). (see 2.3 below for further discussion).

### 1.5 Orientation columns

An orientation column sub-structure will be modeled using a local complex logarithmic mapping and a unidirectional lateral inhibitory operator ( Schwartz 1977c ) ( Schwartz 1984q)

### 1.6 Radon transform and orientation columns

An alternate hypercolumn structure will be examined by using the several operators (e.g. backprojection, hilbert transform, directional derivative) associated with the Radon transformation.

## 2. Applications of cortical architectures to image processing

### 2.1 Data compression

One major and obvious computational use of the space variant architecture of human visual cortex is data compression. The first goal of this study is to use the simulation tools developed above to study the nature, utility, and magnitude of the data compression afforded by the complex logarithmic structure of the human visual field.

### 2.2 Visual invariances

The complex logarithm has a well known size and rotation scaling property associated with its structure. For a given fixation point, movement towards or away from the image, or rotation about the fixation point, results in a linear shift of an invariant image. For non-fixed motions, on the other hand, simple translation of the image results in a change in size and gross shape of the image details, but not in local angular relationships ( conformal property of the complex logarithm). Furthermore, a more complicated form of projective invariance ( Schwartz 1980c) is also associated with the complex logarithmic mapping. An examination of the utility of these various aspects of perceptual invariance will be made, using the modeling tools developed in the first part of this proposal.

### 2.3 Columnar architecture, ocular dominance columns, and difference mapping.

A simple algorithm for stereo segmentation, which does not depend on the usual ( very expensive) approach of matching corresponding points ( Schwartz 1982 ), but which is immediately provided by the columnar architecture ( ocular dominance columns) of striate cortex, will be investigated with real images.

#### 2.4 Columnar architecture, orientation columns, and shape encoding.

The periodic mapping of orientation to "orientation columns" in striate cortex suggests that some form of periodic boundary orientation analysis, such as the method of Fourier Descriptors, is utilized by the human visual system ( Schwartz 1980a) This suggestion has been experimentally supported by data obtained from later cortical processing stages ( infero-temporal cortex) ( Schwartz et al, 1983 ) The computational utility of a periodic orientation map is to be investigated in the present study by using the modeling tools of part 1 of this proposal to construct images mapped via such an architecture, and then studying the utility of this architecture for shape encoding and image description.

#### B. Progress during previous funding period 4/15/83 - 4/15/85.

##### 1. Computer simulation, graphics, and imaging processing environment.

During the past funding period, an environment for simulating cortical architectures, and for studying the possible computational significance of these architectures for image processing, has been constructed. During this period, the P.I. (Eric Schwartz) was appointed as visiting Professor of Computer Science at the Courant Institute of Mathematical Sciences. This appointment allowed for the setting up of the current AFOSR equipment within the space of the Experimental Computer Sciences and Robotics Laboratory of the Courant Institute ( located at the Wash. Square Campus of NYU).

This provided a good opportunity for us, since we were able to make use of many of the facilities of the Courant Institute.

##### 1.0 Computer Hardware environment

Our proposal specified acquiring a Charles River Data Systems M68010 micro-processor for image processing and

computer simulation work. Instead, we acquired a SUN Microsystems M68010 system. In retrospect, this choice was a correct one. The SUN environment is essentially identical, both in use and performance, to the VAX-750 Berkely 4.2 Unix environment. This allowed us to make use of the large body of existing software in the Berkely 4.2 environment, which is the standard academic computing UNIX environment. Thus, we were able to make rapid progress, once we had received shipment and installed our equipment. Our initial orders were placed within one month (June 1983) of the start of funding, we received our equipment in November of 1983, and had set up our basic computer environment (i.e. a SUN workstation, 150 megabyte disk, Berkeley Unix, etc.) within the next month.

For image processing, we are using Image Technology Incorporated (ITI) frame buffers ( 3 buffers) with an RGB analog board ( 8 bits in, 8 bits out to 3 8 bit RGB lookup tables). In addition, this equipment has an image array processor ( ALU) which allows us to do rapid image processing computations, such as convolution, etc. The SUN, together with these ITI imaging boards, forms the basis of our ability to model cortical architectures and algorithms for image processing. However, we first needed to construct a library of basic device driver and applications programs in order to use this hardware, as will be next described in the next section.

#### 1.1 Computer software environment

##### a.) Image processing support: low level

In order to make use of our system, it was necessary to construct a basic library of device drivers, as none was available for the ITI/SUN environment. This was a major, and unavoidable task. We currently have written approximately 150 subroutines in order to have full use of this system for image processing. Although this is an ongoing task, as we add routines to this library as the need comes up, it was largely completed in late spring of 1984.

##### b.) Image processing support: high level

Given a library of routines to control our image processing equipment, we also needed to obtain high level routines to perform standard image processing operations. In this case, we were able to obtain a library of image processing routines, which was written in "C", and which required only minor work to customize to our hardware. This project was also completed in late spring of 1984. Most of the illustrations in this report were prepared with this library of ITI device drivers, support routines, and image processing routines.



c.) Solids modeling: high level

In order to obtain model images, we originally proposed to use both "real world" stimuli from photographs or videotape, synthetic images generated by computer, and images generated by photography or video of "model environments".

In order to provide a capability for generating synthetic images, we have ported a standard solids modeling program, movie.BYU, to our system, and written the necessary device drivers for our hardware. Figures 1 and 2 shows examples of solid model generation capability of this system.

d.) Model environment

In practice, we found it very difficult to obtain images from the "real world" under controlled conditions. Changing light, pedestrians, cars, and general lack of control of image variables such as scale, disparity, motion, color, etc. provided arguments that a model environment would be useful. This was set up using "HO Gauge" models, providing a scale of 1:86. Figure 3 shows a stereo pair obtained from this environment. 4 shows an example of a motion image obtained via photography and frame grabbing. Our experience to date suggests that this model approach is an inexpensive and convenient means of obtaining reasonably realistic images under controlled conditions, with respect such image variables as stereo disparity, parallax, light gradients, color, velocity, etc.

e.) Image data acquisition

The ITI frame grabber hardware/software system described above provides the basic means of entering image data into our system. Since our applications involve stereo, color, motion, as well as high-resolution still frames, some attention needed to be paid to the means of entering image data. A convenient solution to these problems was found by using a conventional 3 color photographic enlarger, together with a CCD camera, to acquire image data. The CCD (charge coupled device) camera is mounted on an X-Y micro-manipulator stage. This allows us to precisely position successive frames with respect to one another, as well as to obtain large image "mosaics". These mosaics are obtained by using the enlarger to project an image that is larger than the area of the CCD chip (e.g. by four times). Then, frames are "grabbed" and the X-Y manipulator is used to move the CCD chip to the next mosaic position. The accuracy of the X-Y manipulator, together with the intrinsic geometric stability of the CCD camera, allows large mosaic's to be built up from individual 512x512 frames. This is important in our simulation of the log map structure of human visual cortex, since extremely high resolutions are needed in the "foveal" area (i.e. order of magnitude of

minutes of arc), while large areas of visual space (e.g. 10-40 degrees) need to be covered.

It is also possible to obtain full color images with this system. The basic hardware (ITI) only supports 8 bits of "frame grab". Full color images require on the order of 15-24 bits (5-8 bits for each of the three primary colors (RGB)). The color enlarger which we are using has separate controls for RGB. Therefore, by illuminating a conventional color slide or negative, and successively grabbing the red, green and blue components of the image, we obtain a full color (24 bit) image. Figure 5 shows an example of three RGB frames acquired in this way. This capability allows us to pursue the proposed studies of color segmentation, and other aspects of color vision.

The ability to acquire "motion" or velocity data, is also solved by this system. The X-Y manipulators allow precise registration of successive frames to be accomplished (The estimated accuracy of this positioning is on the order of 10-20 microns, or one pixel element of the CCD camera). In addition, we constructed a jig which allows pin registered 8mm movie film to be passed through our enlarger/CCD system. Thus, extensive motion sequences may be dealt with by this equipment.

#### f.) Color and motion image display

Since we are able to acquire color and motion image data, it is convenient to be able to display these images. Motion data can be displayed three frames at a time: we grab three successive frames of a motion sequence, and then toggle our three frame buffers at the maximum frame rate of 30 frames/sec. This allows us to get a fair preview of our motion data. Displaying full color images was more difficult. Our equipment has full RGB output, but only 8 bits of input may be fed into this system, through lookup tables. This is typical of most low to moderate cost graphics and image processing equipment. Since our full color images are 24 bits deep, there is a problem in choosing how to truncate this color data, and yet still obtain a satisfactory color display. We solved this problem by using a histogram equalization technique based on the space filling "Peano" curve (Lehar et al, 1983). An example of a 3D Peano curve is shown in figure 5. This (one dimensional) curve fills space: by passing through a 3D color space (axes R-G-B), and histogramming the color values, it is possible to determine the 256 most "popular" color values out of the  $2^{24}$  possible color combinations. Then, using the available lookup tables to display these 256 "best" colors, surprisingly good quality full color images may be obtained. Figure 5 shows an example of such a full color image, which corresponds to the three separate RGB frames shown in figure 5 above.

## 1.2 Global topographic mapping of visual cortex

The mapping of the visual field ( or surface of the retina) to the surface of layer IV of primate visual cortex is well approximated by a map function of the form  $\log(z+a)$ , where  $z$  is a complex variable representing position in the tangent plane to the visual field, and where "a" is a real constant which must be fit to the data.

This map function, proposed seven years ago, ( Schwartz 1977 ) has had a controversial history. Although at the present time there is no other full two dimensional map function proposed in the literature, there seems to be a reluctance to consider mappings of the form  $\log(z+a)$  as a descriptor of cortical topography. However, recently Tootel et al ( Tootell 1982 ) have provided strong experimental support for this map function. They presented an actual 2-deoxyglucose "picture" of this mapping. This data is in surprisingly good agreement with previously published mappings of the form  $\log(z+a)$ , with the constant "a" equal to 0.3 degrees. ( Schwartz 1984 ) Figure 6 shows a superposition of the Tootel et al(1982) data and a previously published map estimate ( and one which was submitted in our previous AFOSR proposal, before the publication of Tootels data). Although the fine details of cortical topography are still begin debated, it seems to be self evident from this figure that a map of the form  $\log(z+a)$  provides a useful first approximation.

The first version of our mapping program uses table lookup to provide the log map. This provided a fast means of providing the log transform on large image files. In this method, we calculate a large table, which specifies the mapping of each point in the image (retinal) plane to each point in the cortical plane. This "lookup table" can then be used on all subsequent images to effect the log mapping without the large overhead of calculating .25 Meg worth of complex log functions. For a given fixation point and image size, one lookup table will service all images. Using this method, it takes about 20 seconds to perform a log transform on a 512x512 image ( .25 Meg pixels), compared to about 10 minutes by direct calculation.

The second computational aspect of the complex log mapping is related to variable spatial resolution. Near the center of the image (i.e. the fixation point), there is a one-to-one mapping of image (retinal) pixels to complex log mapped (cortical) pixels. As one moves away from this central fixation point, the mapping becomes many-to-one: one (cortical) complex log point receives contributions from many image points.

Our first version software effects this many-to-one mapping by simple averaging. All image points in the

(receptive) field of a single cortical point are algebraically averaged. Once again, this gives a fast algorithm: in fact, about fifteen of the twenty seconds which it takes to perform this mapping are due to this averaging operation.

Examples of this mapping software, applied to an image of the space shuttle and to an image of a model tank, are shown in figures 7-15, for several different values of the constant "a" in the function  $\log(z+a)$ , and for several different fields of view. Also shown are the "retinal" versions of some of these mappings (figures 8,12 and 13).

The "blocky" quality of these image is due to the use of integer arithmetic in our mapping and averaging programs. Our intention in this first level software was to obtain a map function which executes rapidly, in order to allow the study of many different kinds of images. The log mapping actually maps pixel locations into non-integral locations. In order to treat this "smoothly", we would need to use bilinear values in our mapping, and also to "average" by floating point interpolation, rather than the present simple integer averaging. We plan to implement this "floating point" version in the near future, but also expect it to be much slower than our current version. The "blockiness" of the images is at present mainly a cosmetic problem, and we are willing to accept it since it gives us a very fast program which is convenient for initial exploration.

The "retinal" images shown in figures 8,12 and 13 are interesting, since they provide an accurate estimate of the spatial information provided by single scans of an image, to the human visual system. This display reinforces two major points:

1. There is a tremendous data compression provided by the topographic structure of the human visual system. Figure 8 provides a graphic illustration of this data compression by showing an image of the "cortical", "retinal", and "world" version of an image. The tiny size of the cortical image is due to the fact that it is shown to "scale". The data compression relative to the "world" or original scene, is about 100:1 for this particular choice of "a" and visual field extent.

2. Any single fixation provides high resolution data from a very limited region of visual field. Therefore, algorithms for eye scanning, and computational procedures for "blending" together different scans, are of paramount interest and importance.

These two points, related to data compression and eye scan paths, will be discussed in greater detail below. At present, we will continue presenting our progress report.

### 1.3 Space variant digital filtering.

Spatial processing which occurs at the retinal level is usually modeled by a circularly symmetric spatial filter. This filter has been modeled, at various times, by lateral inhibition, by using difference of gaussian filters, by the use of the Laplacian operator ( Kovacsny and Josephs, 1958 ), or by one of many possible symmetric digital filters. These methods are, to our understanding, largely equivalent.

Less attention has been paid to the fact that the size of the filter used must scale correctly with cortical magnification factor. Additionally, it would be highly desirable to provide a multiple resolution spatial filter, which itself scales correctly with cortical magnification. The multiple resolution aspect of this filter is interesting because of the very large amount of physiological and psychophysical evidence which suggests that multiple scales of analysis occur at the early levels of visual processing.

An efficient way to provide both multiple resolution at a point, and spatial filtering which is circularly symmetric and which scales correctly with cortical magnification factor, has been implemented as follows.

First, we construct a Laplacian Pyramid. This data structure has been applied by Burt and Addelson ( Burt Addelson 1983 ) to a variety of image processing applications. The basic idea of the Laplacian Pyramid is to essentially construct difference of Gaussian filtered images at successively coarse levels. First, the original image is sub-sampled to a smaller scale, smoothed with an (effectively) broader Gaussian, expanded to the original scale and subtracted from the original image. This produces an image which has been filtered by a "difference of Gaussian" filter. This represents a discrete approximation to applying the Laplacian operator to the image ( Hummel 1984 ). Figure 14 shows an image of the space shuttle, with the various levels of the Gaussian Pyramid ( i.e. smaller and smaller and smoother and smoother images ).

The utility of this pyramid data structure is that it provides a model for an image which has been processed by "retinal" circularly symmetric receptive fields, at multiple scales of spatial resolution at each point of the image.

Burt's scheme for computing the pyramid is attractive, since it only uses  $4/3$  the amount of space and computer time to calculate the entire pyramid (multiple resolution) as it does to compute the top level. Thus, if one is going to bother to do any spatial filtering at all, one may as well do it at multiple scales!

The pyramid, as described above, and illustrated in figures 14 and 15 , provides us with multiple scales at each point of the image. However, in addition, we need to

construct a space variant resolution, due to the logarithmic structure of the primate visual system.

There are two ways to do this: the first way might be to work at the "retinal" end, and use a filter whose width increased appropriately with retinal eccentricity. This would be expensive in computer resources. The second way is to use a constant size filter at the "cortical" end, and to then use the inverse log mapping function (i.e. the complex exponential) to create the retinal image. This is the method we have chosen. It is highly efficient, since we already have the logarithmic point mapping implemented. Thus, in practice, we merely have to compute the pyramid function at the cortex, and then back-map the pyramid to the "retina", using the same table used for the log mapping. This is extremely fast, since the cortical image is very small ( see figure 8), so that we gain full benefit of the natural data compression of the log mapping in this application.

We view the data structure of the Complex Logarithmic Laplacian Pyramid as the basic representation of the large scale spatial structure of the primate visual system. It incorporates the effects of circularly symmetric lateral inhibition in the retina with the space variant nature of the global topographic mapping.

#### 1.4 Ocular dominance columns

Ocular dominance columns are the thin patches (0.5 mm) of cortical tissue which received alternating input from the left and right eyes. They look vaguely like "zebra stripes" ( see figure 6 for Levays tangential reconstruction). We are examining the possibility that a form of textural encoding of binocular stereo data is provided by this system ( Schwartz 1982 ). Figure 16 shows a computer enhanced reconstruction of a section of striate cortex, with the ocular dominance columns from one eye enhanced by cytochrome oxidase.

A major goal of ours is to reconstruct the spatial configuration of two images (left and right), at the level of striate cortex, due to both cortical topography and ocular dominance column interlacing. We provided hand-drawn estimates of the back-mapping of ocular dominance columns to the visual field ( Schwartz 1977 ), and Hubel and Freeman have also attempted to provide this trajectory mapping, but used an apparently incorrect map function, based on a map of the form  $\log(r+1.7)$ , with radial symmetric mapping. It is very unclear what this map function would look like: Hubel and Freeman did not present it, although they used it in their work. Our studies have indicated that map functions such as this are in error by at least 100%: they fail to even come close to fitting the opercular surface of the cortex. However, since the map function  $\log(z+0.3)$  does fit nicely over the opercular surface, it is relatively simple to generate

not only the ocular dominance column trajectories, but an actual image as well. Figures 17,18, and 19 illustrate our initial attempts at this work. We feel that it is very important to use higher resolution images: we can only map the central 2 degrees at present, and the detail is not sufficient (due to using 512x512 sized images) to really study this central area. Nevertheless, figures 17,18 and 19 do represent, to our knowledge, the first correct presentation of both ocular dominance column trajectories, and of a stereo image encoded at the level of striate cortex.

#### 1.5 Orientation columns simulation.

In order to have a complete model of striate cortex architecture, the orientation column pattern described by Hubel and Weisel must be constructed. There is a slight problem here: although the global topography and ocular dominance column patterns are reasonably well understood at the present time, the orientation column (or hypercolumn) description has undergone some dramatic revisions quite recently. It's detailed structure is still uncertain.

Specifically, the original model of Hubel and Weisel (Hubel and Weisel, 1974) consisted of parallel "slabs" of cells, all tuned to the same orientation. Each slab was roughly 50 microns wide, and there were roughly 10 slabs spanning the full 180 degree set of orientations in a "hypercolumn".

Recently, Hubel and Livingstone (1984) have noted, by using the cytochrome oxidase method, that in the center of each ocular dominance column, there is a "hole" in the orientation tuning pattern. These so-called cytochrome oxidase spots have a higher than average spontaneous (and metabolic) activity (so they stand out in cytochrome oxidase stain, which is a metabolic marker). Furthermore, there is little or no "orientation tuning" inside the area of these spots, which are most prominent in layer IV of striate cortex, but which extend into the upper layers as well.

The discovery of this phenomenon was a welcome result to our lab. This is because a model presented for the local structure of the cortical hypercolumn, seven years ago, had a rather prominent "hole" in the center (Schwartz, 1977) due to the singularity of the complex logarithmic mapping. At this time, this was a difficulty of our model: there was no evidence of a lack of orientation tuning in the center of cortical hypercolumns, nor there was there any evidence that hypercolumns were "modular", as opposed to "continuous" runs of orientation change. Thus, our earlier model both provides a simple account of the existence of roughly parallel "orientation columns", but also essentially predicted the existence of "holes" of orientation tuning (see figure 21

and 22) in the data recently published by Livingstone and Hubel.

We have implemented a preliminary version of the local complex log mapping. The method was to use the same "table lookup" program discussed earlier in the context of global topography, applied to a small patch of image. In addition, a uni-directional lateral inhibitory operator is applied to this mapping. The details of this series of log mapping and uni-directional inhibitory operator are discussed in earlier work. ( Schwartz 1977c )

Perhaps the best way to illustrate this idea is by the sequence of images which are produced.

In figure 20, we have a log mapping, applied to a field of parallel bars. The map function is  $\log(z+1)$ , and the size of the field is 9 on either side of the center. (Note that the units 1 and 9 above are now in local units of distance, rather than in units of visual field angle for the global mapping.)

It can be seen that a specific pattern of afferents is constructed, but that the mapping of the "stimulus" lines is diffusely spread all over the area of the "hypercolumn". There is little evidence of orientation banding from the afferent pattern alone.

In the next column of figure 20, we apply a uni-directional high pass filter. This is meant to simulate the effects of a directional lateral inhibition which is oriented in a direction parallel to the ocular dominance column borders, as specified in the original model. ( Schwartz 1977c )

Note that this directional derivative leaves a "banded" pattern which is broad, but localized to a segment of the hypercolumn.

In the third column of figure 20, a smoothing operator ( medial filter ) is applied, and in the fourth column, a threshold is applied to simulate neural firing.

The result of these operations is four rather broad "orientation" columns, with a prominent "hot spot" or "hole" in the center of the hypercolumn. This central area is caused by the convergence of all orientations there: this is like a local "fovea", in the sense that it is the origin of the local log mapping.

Thus, figure 20, which is a simple simulation of a local log mapping model of cortical hypercolumns, captures much of the currently understood detail of this feature of cortical architecture. "Columns" of tissue will be responsive to all lines of a given orientation. Rotation of these



lines will cause a linear shift of the center of this slab, or orientation column, as is indicated in figure 21. Figure 22A shows this simulation for an array of columns, with two different orientations superimposed and thresholded. The "hot spot" in the center of this simulation is clear. There is a failure of orientation tuning in the center of the column, together with an area of increase "spontaneous" activity. In other words, the current state of Hubel's model of the cortical hypercolumn is captured in this simulation, and in effect, was predicted by this model roughly seven years before the data had been recorded.

The modeling demonstrated in figures 20-22 is still in a fairly preliminary stage. Nevertheless, it is a quite fast computer implementation, and should be easily applied to actual image data (rather than to patterns of parallel bars) in the very near future. The principal question which remains concerns the nature of the spatial frequency tuning in the center of the C.O. spots. Some very sketchy reports suggest that the spatial frequency tuning in these spots is of lower frequency than that in the orientation tuned areas. (Tootell 1981) This may be because there is less inhibitory drive in the spot areas; this would explain their higher spontaneous activity, and circularly symmetric fields. On the other hand, the simple model presented above also shows higher "spontaneous activity" in the central spot regions, due to the convergence of many more inputs in those regions. Furthermore, the use of the term "spatial frequency" to compare edge responses to circular responses, as done by Tootell et al (1981) is questionable. An edge has a long region and a narrow region: the long region is presumably due to excitatory summation, while the narrow region is due presumably to lateral inhibition. Looked at in this way, the central C.O. spot cells may have much smaller "excitatory" summation areas, as would be expected from the complex log model.

#### 1.6 Radon transformation model of cortical hypercolumns.

We are also exploring an alternate model of hypercolumn simulation, based on the Radon transformation (Schwartz 1984).

The Radon transformation describes how an image may be reconstructed from "line integrals". It is the basis of computed tomography, where the line integrals are, e.g. x-ray attenuations, as in CAT scanning.

In the present context, the Radon transform indicates how one may go from bar-like summations, to a complete description of a scene.

Our current state of progress on implementing this work is as follows: we have obtained a set of programs from

Donner Labs at U.C. Berkely, which perform many different aspects of the Radon, inverse Radon, and associated transforms. We have not yet implemented any of these programs to date. It is expected that in the next six months, we will have begun using the Radon modeling tools to study this alternate model of hypercolumn architecture.

## 2 Applications of cortical architectures to image processing

### 2.1 Data compression.

A major computational problem for visual systems (biological and artificial) is the sheer mass of data which is associated with images. A 512x512 image (moderate to low resolution in the current context) has .25 million pixels. A 4000x4000 image, which more closely approximates the (maximal) spatial resolution of human vision (when spread over a full 80 degree field), has 16 million pixels.

The solution which primate (and cat) visual systems have found is to use a space variant system, coupled with an attentional system which allows selective fixation. Thus, for example, the human visual system has a resolution which is roughly one minute of arc centrally, but which falls off rapidly to about one degree at 10 degrees of view, with further falloff in the extreme periphery.

For a single fixation point, a very high resolution image is provided in a very small part of the visual field (the foveal area), while a very large part of the visual field is still covered at lower resolutions. Since the choice of foveal fixation is attentionally controlled, this form of data compression is quite different from virtually all current approaches (in machine vision). On the one hand, the amount of data compression is enormous. In order to provide a conventional image system with resolution comparable to human vision over a field comparable to the human visual field, one would require  $\sim (80 \text{ degrees} \times 60 \text{ pixels/degree})^2 = 4800^2 \text{ pixels} \approx 23 \text{ million pixels}$ . In order to achieve a comparable temporal resolution, one would have to provide roughly 30 such frames/sec, which is a bandwidth of 690 Megapixels/second! It is fairly obvious that potential high-performance machine vision systems must follow some similar course to that followed by the human visual system in order to reduce the amount of data. Conversely, one can get a rough estimate of the amount of data compression which is provided by the human visual cortex. If one takes the number of hypercolumns as roughly equal to the number of "pixels", i.e. spatial resolution elements across the surface of striate cortex, then there are only on the order of 1-2 thousand "pixels" covering the entire visual field. This is a very rough estimate, but is in agreement with the measurements and calculations of Dow et al (1981),

who find that the extent of a hypercolumn in primate visual cortex is about equal to the corresponding value of visual acuity. This calculation leads to a data compression value of  $23 \times 10^6 / 2 \times 10^3$ , or roughly 10,000:1.

The previous method of calculating data compression due to the space variant nature of primate visual cortex is very rough. The correct value might be an order of magnitude or two lower. Even so, there is an incredible amount of data compression provided by the space variant nature of topographic mapping. However, this data compression is provided at the expense of requiring an attentional system which must almost instantaneously compute the best value for the next fixation, as well as other systems which must integrate space variant views of a scene together.

One of the major goals of the present work is to examine these two questions. We have just begun to study the integration of different fixation point views to produce "composite" or blended views of a scene. Although at an early stage, these efforts demonstrate an ability to construct such "blended" space variant views of scenes. The second aspect of this work, studying attentional models of choosing eye scan paths, is to be studied in future work.

#### 2.1.a Data compression for various values of log mapping constants and fields of view: single fixation point.

We used the log mapping (table lookup) program described to generate both "cortical" and "retinal" views of several scenes, obtained from real objects (space shuttle) and model objects (tank and model town). One area of uncertainty involved the choice of parameters to use to model the topographic mapping of human visual cortex. At the current time, agreement exists that the fall-off of cortical magnification factor is roughly logarithmic, with functional fits that range over functions like  $\log r + 1.7$ ) (Hubel and Freeman, 1977; Drasdo, 1977; Rovamo and Virsu, 1979;) to functions like  $\log(r+1)$  (Tostell et al, 1982) to functions like  $\log(r+0.7)$  (Van Essen et al, 1983),  $\log(r+0.3)$  (Dow et al, 1981). (Hubel and Freeman 1977) (Drasdo 1977) (Rovamo Virsu 1979) (Tostell 1982) (Van Essen 1984) (Dow 1981)

Note that the implicit or explicit map function of all of the above experimental studies is not the complex logarithm, but some unspecified function which has a radially symmetric dependence (for Hubel and Freeman, Van Essen et al, Dow et al), or a radially asymmetric dependence (Drasdo; Rovamo and Virsu) which follows that of the retinal ganglion cells, and therefore has the horizontal meridian longer than the vertical meridian.

In fact, the correct cortical asymmetry is one which causes the representation of the vertical meridian to be about 30% longer than that of the horizontal meridian, for the central 5-8 degrees. This is shown directly in the data of Tootel et al(1982), which is the only (2DG) mapping of striate cortex performed to data. This is also shown directly in figure 6 of this report ( and this point is made directly in a forthcoming report in SCIENCE ( Schwartz 1984 )

In the current studies, we have used both the map function  $\log(z+0.3)$  ( which we feel is the best current estimate, as noted above), and the map function  $\log(z+1.7)$ , which represents the more "traditional" value favored by Hubel and Freeman, Drasdo, and Rovamo and Virsu. The principal difference between these two map functions, which almost certainly span the range in which the "correct" map function is contained, is in the degree of magnification in the foveal region. The map function  $\log(z+0.3)$  (e.g. figure 8) has a much larger degree of magnification than the function  $\log(z+1.7)$  (e.g. upper right corner of figure 9). For example, the ratio in magnification at 0 degrees and 10 degrees for the function  $\log(z+1.7)$  is equal to (along the horizontal meridian)  $11.7/1.7 = 6.8 : 1$ . The corresponding ratio for the map function  $\log(z+0.3)$  is  $10.3/0.3 = 34 : 1$ , while the corresponding ratio for another recent measurement ( Van Essen) of a map function like  $\log(z+0.7)$  is  $10.7/0.7 = 15 : 1$ .

Figures 8 through 15 show log mappings, using various map function parameters (  $\log(z+0.3)$ ,  $\log(z+1.7)$ ), fields of view ranging from 5 degrees to 25 degrees. Also shown in figures 8, 12 and 13 is the "retinal" view, obtained by taking a scene, computing its log mapping, then back-projecting to the retina with the complex exponential.

Shown in figure 8 is a comparison of the log mapping and the original scene. The log mapping is scaled so that nearest neighbor pixels in the foveal representation are exactly nearest neighbors in both the "retina" and the "cortex". This establishes a "scale" for data compression. The actual data compression is given by the ratio of the log map and the "original" scene's areas. In figure, this is roughly 100:1. (Note that the larger values of compression calculated above were for much larger extents of visual field, i.e. out to 80 degrees).

The "retinal" versions of these scenes are quite interesting. These correspond, in a fairly realistic way, to the spatial information which is available to the human visual system in a single fixation. There is an extreme fall-off in detail with eccentricity, as is well known, but which is perhaps made more explicit in these figures. The space-variant nature of the mappings seems to destroy the

detail outside of the immediate vicinity of the foveal region. However, if the viewer stands back from these figures, it is clear that a surprising amount of the "Gestalt" of the figure is still retained in the peripheral representation. In fact, by careful generation of figures like these, it should be possible to create a "retinal" image which, when fixated at the correct point, should be indistinguishable from the full resolution image, since the fall-off in detail may be matched accurately with the similar fall-off in detail of the human visual system.

### 2.1.b Multiple fixation points

One of the most powerful "illusions" of human vision is that of a stable, homogeneous and uniform visual field. In fact, as the previous examples demonstrate, the visual field is highly non-uniform. In fact, any small region of the visual field receives widely different inputs as the eyes scan the scene. Conversely, any given part of the visual scene is imaged with widely varying spatial resolution, as the eyes move across the scene. The issue of whether or not there is "summation" across fixations in order to provide a stable percept of the world has been argued back-and-forth: we would like to avoid this aspect of considering the problem of multiple fixation points. Rather, as it is a matter of everyday experience that we somehow do perform a "merging" of multiple views of a scene, we will concentrate below on methods of representing this synthesized image. Figure 7 shows the "visual scene", which is the space shuttle, and 9 "fixation points", which are marked by crosses. Figure 8 shows a single log mapped view, for a fixation point near the "d" of "United States". Figure 23 shows these 9 cortical views synthesized into a single "minds eye" view representing 9 saccades of scanning the scene of the space shuttle. This issue of multi-fixation point views will be addressed further in this proposal.

### 2.2 Visual Invariances

The complex log mapping has several well known invariance properties associated with its structure. For a given origin (fixation point), size or rotation of the image results in a linear shift of the complex logarithmic image, outside of the immediate area of the foveal projection. This immediate area is effectively defined by the value of the constant "a" in the map function  $\log(z+a)$ . For values of  $z$  less than  $a$ , this map is roughly linear; for values of  $z$  greater than  $a$ , the map is logarithmic and has the invariance properties described above. For a value of  $a = 20$  minutes of arc, as reported by Dow et al (1981), there is a small area of 20 minutes radius which has a roughly linear topography; the remainder of the visual field (at least out to 20-30 degrees) has logarithmic structure.

These invariance properties of the log mapping are well known, and have been the object of a large number of studies in machine vision ( Brousil Smith 1967 ) ( Casasent and Psaltis 1977 ) ( Chaiken and Weiman 1979 ) ( Jain ) ( Shenker 1981 ) and in biological vision ( Schwartz 1977a ) ( Cavanagh 1978 ). These studies may be classified along two major lines. One of these is to combine the log mapping an initial ( and following ) Fourier transform ( e.g. Brousil and Smith, Casasent and Psaltis, Cavanagh ). This has the advantage that the initial fourier transform ( amplitude map ) is shift invariant, so that the resultant map is invariant to shift, size and rotation. Cavanagh (1978) has proposed a model of this sort for human vision. However, this model suffers from one fatal flaw: the initial fourier transform is space invariant. Thus, a system of this kind would have constant magnification, or acuity, across the entire visual field. This is obviously incorrect for human vision, as discussed in ( Schwartz 1981 ) In addition, there is strong evidence against anything like a global fourier transform occurring in the early stages of human vision, ( Schwartz 1981 ) and such a transform is essential for this type of model.

The second class of applications of the complex log mapping is based on using the map properties of the complex log, by itself, as an aid in computing invariances (Chaiken and Weiman, Jain, Schwartz). This is a more difficult ( if more realistic ) undertaking. The complex log distorts distance relations when the image is shifted, although it preserves local angles ( or curvature ) since it is a conformal map. Because curvature is invariant under the log map, perhaps the best candidate for making use of its properties is a combination of some curvature based descriptor, such as Fourier Descriptors, with the spatial properties of the log mapping. This possibility will be discussed further in this proposal. Our initial efforts at studying the symmetry properties of the log mapping have been confined to generating images of scenes which undergo size changes ( or perspective changes ), as they appear in the complex log plane.

In figures 24, 25, and 26 we have used our log mapping software to generate images of a "town" and a "tank", modeled in HO gauge, as a camera is zoomed along its visual axis. The map functions used to generate these figures, and the field of view, of the figures, is indicated in the figure captions for each case. The original unmapped view of each tank is shown in figures 9, 10 and 11.

These figures were generated by photographing a model scene with a camera with a grid reticle, which allowed continued focusing on a single point of the scene ( e.g. the star of the tank ) as the camera was moved. Three frames, at roughly one foot intervals, were recorded. In order to provide some form of velocity scale to this "zooming" motion,

one might consider these frames to be part of a movie or video sequence, at 30 frames/second. The one foot physical movement of the camera is scaled to 86 feet, because HO scale is 86:1. Thus, the velocity implied by these choices of scale is very large: 30\*86 feet/second, or roughly 1600 miles/hour. Slower "velocities", i.e. smaller movements of the camera, would lead to much smaller effects. It was our intention in these initial studies to provide large (gross) velocity effects.

Several results can be seen in these frames.

1. Projective scaling.

For all stimuli in the log-mapped frame outside of the central foveal projection (e.g. either 20 minutes of arc, or 1.7 degrees, depending on the map function used, there is a conversion of size change and positional shift (due to "zooming") to a simple linear shift, parallel to the horizontal axis.

2. Projective scaling for non-visual axis motion.

One problem of using the log mapping in image processing applications is the special role provided by the fixation point. The symmetry properties of the log mapping only strictly hold for motions about the fixation point. Thus, if an observer (camera) moves in one direction, but "looks" in another direction, the symmetry properties of the log mapping are compromised. In figures 27, 28, 29, and 30 we have begun to study this problem by considering off visual axis motions. In this example, the fixation point is moved successively 1, 2, 3, 4 and 5 degrees from the "impact" point, which is the star on the tank (for the largest image of the series). At the present time, we have not yet evaluated the analysis of "zooming" motions for which the impact point and fixation point are not the same, as shown in these figures. It does seem to us that there is still a substantial similarity in size and shape of the peripheral image detail, even though it is not strictly identical as in the case of motion along the optical axis (as in figure 31 for example). Thus, the differences in the retinal projections of the stimuli, under this motion are very large (e.g. 100% differences in size). The cortical scenes, for corresponding size differences, seem to be much more similar. If so, then the disadvantages of the log mapping for motions which are not along the line of sight may be relatively minor. We are currently considering this issue, and attempting to develop ways in which to quantitate these effects.

The examples shown above are preliminary efforts at studying the invariance properties of the complex log mapping. It is clear that these invariance properties exist, and that the human visual system (or machine systems) might conceivably utilize them. Our efforts in the immediate future will center around generating a much larger variety of image conditions, and, developing some way of quantitatively evaluating whether these symmetry properties are of

importance ( or, equivalently) of how to use them. This is a difficult problem: it is evident from the figures shown above that there seems to be a surprisingly good invariance to "zooming", even in the unfavorable cases of very large simulated velocities, and for small off-axis components of motion. In order to pursue the question of the utility of these properties, we need to provide a segmentation and classification scheme: the comparison of the performance of this scheme, using the complex log, compared to conventional schemes, seems to be the only way to determine the utility of the complex log mapping as an aid in image analysis. Proposed methods of performing these evaluations are outlined below, in our proposed future work.

### 2.3 Ocular dominance columns and difference mapping.

One possible computational use for columnar architecture, i.e. the interlacing of two slightly different aspects of a visual stimulus in adjacent "strips" of cortical tissue, is the extraction of difference mapping by spatial filtering ( Schwartz 1982 ). Examples of this are: 1.) ocular dominance columns: slightly different left and right eye views are interlaced as columns in visual cortex. The difference between these views is a cue to stereo disparity. 2.) color columns: slightly different chromatic views may (experimental evidence premature) be interlaced. The difference of these maps are R-G or B-Y maps, which may be cues to color segmentation.

We have generated initial versions of the ocular dominance column map of two slightly different visual scenes ( i.e. the space shuttle). The map function  $\log(z+0.3)$  is an excellent fit to the shape of the operculum ocular dominance column reconstruction of Levay et al (1975) (see figure 6). We have used this map function, together with Levays ( Levay 1975 ) ocular dominance column map, to generate an interlaced cortical view of a binocular disparity stimulus (figure 17). Figure 17 shows the cortical interlacing of a small part of the space shuttle. Figure 18 shows the "retinal" version. This figure may be high-pass filtered (via the Sobel operator, for example), and then mapped back to the retina, as shown in figure 19 . Although still in a preliminary stage, this work indicates the existence of a rough "textural" cue for segmentation via disparity and difference mapping, due to the ocular dominance column interlacing. We plan to continue this work in the next several months by using more realistic images ( from our model environment), more complex images than the space shuttle, and by blending multiple fixation points.



## PSYCHOPHYSICS OF BOUNDARY CURVATURE

INTRODUCTION

A quantitative approach to the perception of contours and shape recognition by the human observer has been proposed which would employ Fourier Descriptors (FDs) of boundary curvature. The periodic orientation analysis performed by the FD method is suggestive of the periodic orientation column pattern of striate cortex. In our earlier work, we proposed that extra-striate visual areas might be sensitive to periodic patterns of boundary orientation (Schwartz, 1980a, 1983, 1984a). This hypothesis received support from studies showing that neurons in the inferotemporal cortex of the macaque were selective for stimuli of specific boundary curvature, as tested by means of FDs (Schwartz et al, 1983).

In the current proposal, we have begun to assess the relevance of FDs to human shape recognition. The questions we have asked are the following:

(1) Does the human observer respond systematically to single FDs? If so, are the responses invariant to changes in size and translation?

(2) How do individual FDs interact in the perceptual domain? Are they susceptible to adaptation? If this is the case, are the characteristics of adaptation with FDs more compatible with an interpretation based on neural fatigue, or do they suggest the operation of active processes of inhibition and/or facilitation? Is there interocular transfer of adaptation, which would indicate a central, as opposed to a retinal, locus for FD processing?

(3) Are the response characteristics to single FDs relevant to the perception of real shapes (in the form of simple closed outlines to which the FD approach is applicable)?

Our findings to date may be summarized as follows:

(1) We have established the regularity of responsiveness to FD stimuli by obtaining the threshold amplitude required to discriminate an FD from a circle (an FD of constant curvature), for a series of FDs at frequencies ranging from 2 - 32. Reverse contrast ratios for luminance produced identical results. Doubling the visual angle (from 1.3 to 2.6 degrees) produced overlapping curves, indicating a form of invariance to size. We plan to test later for responses to shifts in retinal position of FD patterns. These results parallel our physiological observations: in the macaque's inferotemporal cortex, a relatively large number of cells which were selectively tuned to FD frequency also showed an

invariant response to changes in size, luminance, and contrast of FD stimuli (Schwartz et al, 1983).

(2) Interactions among FDs are easily made apparent by means of the adaptation procedure. Discrimination thresholds can be raised or lowered, depending upon both the amplitude and frequency of adapting FDs. At a given frequency, higher amplitudes tend to produce greater impairment of discrimination. With regard to the question of frequency specificity, the higher adapting frequencies tested are both more susceptible to being adapted, as well as being more influential in affecting responsiveness at other frequencies. Further details on this question are presently being examined. Preliminary studies have provided evidence of interocular transfer of adaptation effects.

The most dramatic effect of adaptation with FDs is its prolonged time course. Impaired discrimination has been regularly observed for 15-20 min after the termination of a series of 50 500-msec adaptation stimuli (administered in conjunction with an equal number of discrimination trials).

Preliminary findings show that the adapting effect can be counteracted by the interpolation, prior to post-adaptation tests, of a series of FDs different from the adapting FD. It is unlikely, therefore, that neural fatigue is the operative mechanism. The characteristics of FDs which can successfully counter the adaptation effect are likely to provide clues as to how curvature analysis is performed by the human visual system.

(3) In future work, we also propose to examine the relevance of individual FDs to real shapes (plane closed curves) by using the properties we have already ascertained for individual FDs, their ability to produce selective adaptation and facilitation. We will also attempt to relate the similarity of the FD constituents of a real outline to shapes perceived as similar by the human observer.

The nature of the interactions we have observed among FDs, the types of distortions reported by our observers, are reminiscent of reports on figural aftereffects (e.g., angular displacements, the Gibsonian curvature aftereffects). A series of models have been proposed ( see Oyama, 1978 ) to account for the vast number of observations on figural aftereffects, many of which deal with the interaction or influence of curved contours. Both the models and descriptions of the patterns used as stimuli lack a formal and generalizable descriptive tool for the most salient property of the stimuli used, namely, their curvature. The FD method of boundary curvature provides for that lack, and may furnish the quantitative approach needed to integrate the research on figural aftereffects.

A description of the methods used to address the above questions will be followed by a summary of the results obtained to date. A research plan for future studies will then be detailed.

Psychophysics Abstract

A quantitative approach to shape recognition by the human observer was proposed which would employ Fourier Descriptors (FDs) of boundary curvature. The periodic orientation analysis performed by the FD method is suggestive of the periodic orientation column pattern of striate cortex. Schwartz (1980) proposed that extra-striate visual areas might be sensitive to periodic patterns of boundary orientation. This hypothesis received support from studies showing that neurons in the inferotemporal cortex of the macaque were selective for stimuli of specific boundary curvature, as tested by means of FDs (Schwartz et al, 1983).

The relevance of FDs to human shape recognition can be approached in several ways, and the questions we have posed, for this series of studies, are the following:

(1) Does the human observer respond systematically to single FDs? If so, are the responses invariant to changes in size and translation?

(2) Are FDs susceptible to adaptation? What is the duration of adaptation following exposure to brief adapting patterns? Most significantly, is adaptation to FDs frequency specific?

(3) Is there interocular transfer of adaptation with FDs?

We have established the following:

(1) a systematic relationship between FD amplitude at threshold and FD frequency (an amplitude threshold or sensitivity function) can be generated (Alter & Schwartz, 1984, and manuscript in preparation). This was accomplished by obtaining the threshold amplitude required to discriminate an FD from a circle (an FD of constant curvature), for a series of FDs at frequencies ranging from 2 - 32. Doubling the visual angle (from 1.3 to 2.6 degrees) produced overlapping curves, indicating an invariant response to size. We plan to test later for responses to shifts in retinal position of FD patterns.

(2) Interactions among FDs are easily made apparent with the adaptation procedure. Discrimination thresholds can be raised or lowered, depending upon both the amplitude and frequency of adapting FDs. At a given adapting frequency, lower amplitudes tend to produce a facilitating effect, while higher amplitudes impair discrimination.

Frequency specificity curves can be generated with FD stimuli. They were obtained for 3 adapting frequencies, 4, 8, and 12, in two observers. This frequency selectivity for FDs implies the existence of "channels" for the analysis of curvature.

(3) Interocular effects are present with FD adaptation and are typical of the results reported for the interocular transfer of visual aftereffects.

These findings imply that a periodic analysis of boundary orientation is performed by the human visual system. Furthermore, the data on interocular transfer suggest that this analysis takes place in a locus more central than the retina.

A description of the methods used to address the above questions will be followed by an elaboration of the experimental findings.

## METHODS

### A. Visual Display

FD stimuli were computed, and displayed on a black and white TV monitor. A four-bit 512x512 frame buffer and a PDP-11/23 host were used for computation and display.

The 512x512 resolution of the display was insufficient to display high quality stimuli. Aliasing, or "staircasing" was slight, but clearly visible on our stimuli, which were nearly circular (i.e., small amplitude FDs). In order to avoid this aliasing effect, a "smoothing" or anti-aliasing procedure was used. One way to do this is to convolve a Gaussian kernel with the display. This effectively removes the aliasing. However, it is very slow, particularly on a PDP-11/23. Therefore, we implemented a table-lookup based anti-aliasing routine, which was based on distance of image pixels to the vector segments making up the boundary of the FD. Special care was needed at the end-points of

the vectors, since the near circular nature of our contours required smoothness at the intersection of each vector, as well as in the perpendicular neighbourhood. Our anti-aliasing approach successfully removed the "staircasing", as judged by our inability to see its effects at the stimulus contrast levels employed.

Sets of FDs at different frequencies, amplitudes, phase, size, and luminance were generated prior to testing. Illustrations of FD patterns are provided in Fig. 33. FDs were normalized to equate for area, which also resulted in an approximately equal perimeter for the smaller amplitudes. The average luminance of the entire display was approximately 11 cd/m<sup>2</sup>. The contrast value between target and background was at -.37 initially (dark outline on light background), but was changed to +.71 (lighter pattern than background) to avoid the light-adapting effect of a bright screen. The framed screen subtended 2.75 degrees (height) X 3.33 degrees (length) of visual angle. The visual angle subtended by the FD stimuli was 1.33 degrees for initial experiments, and 1 degree for the experiments on frequency specificity and interocular transfer. The centered fixation spot subtended less than 10 min of arc. The observer viewed the display binocularly, with a viewing distance of 240 cm.

## B. Subjects

### (1) Experiments on thresholds.

Four individuals (one female and three males ranging in age from 17 to 43) participated extensively in these experiments. All observers possessed normal or corrected-to-normal acuity. Observers with optical corrections used them during the experiments. Two observers were naive with respect to the purposes of these experiments. All were provided practice in making the required discriminations prior to testing.

### (2) Experiments on adaptation and frequency specificity.

Four observers participated in this series. Two had undergone the threshold series. Two additional observers were recruited, with the same characteristics as described above.

### (3) Experiments on interocular transfer.

Since this was a group study, 7 new observers were employed. They had been provided extensive practice in a series of binocular FD studies not described in this report.

### C. Procedure

#### 1. First Experimental Series: Tuning Curves

Using the method of constant stimuli, FD amplitudes at threshold were obtained for frequencies ranging from 2 through 32 by means of a two-interval forced-choice procedure. A single session, which lasted approximately one hour, consisted in establishing the threshold for a single FD frequency. Five or more blocks of trials were run, corresponding to the 5-7 amplitudes preselected on the basis of pilot studies, and administered in random order. Each block of trials at a given amplitude consisted of 50 trials, each trial containing one standard and one comparison stimulus, the order of these being randomized over trials as well. The standard stimulus was always a circle, i.e., an FD of constant curvature. The comparison stimulus was an FD at the predetermined amplitude for that block of trials. Stimulus duration was 50 ms, and the ISI was 500 ms. The intertrial interval was approximately 4 secs, including response time. Each block of 50 trials thus consumed 4 mins. A 5-min pause intervened between blocks.

Each subject was adapted to the dim ambient illumination of the testing room for approximately 10 min. Instructions consisted in informing the subject that every trial would contain two stimuli, a circle and a figure similar to a circle but one which would have a number of indentations or lobes. The subject was to specify which interval, the first or second, contained the non-circle, even if the response was a guess. A warning tone preceded every trial by 500 msec, and feedback (a double tone) was provided when the response was correct. Although initial experiments did not include such feedback, repetition of the experiments showed no significant difference in the results. Nine frequencies were tested in this manner, usually on separate days.

#### 2. Second Experimental Series: Adaptation and Frequency Specificity

(1) Adaptation studies using the method of constant stimuli.

Pilot studies using a mid-range frequency had shown a marked adaptation effect on two subjects using only a moderately high adaptation amplitude. In order to determine optimal adaptation amplitudes, as well as to examine the characteristics of adaptation on test frequencies both identical to and one-half octave below the adaptation frequency, the following procedure was

employed on 2 subjects.

Adaptation amplitudes were increased in steps of one octave for frequency 8 (our mid-range frequency), and the test frequencies were 8 and 6, tested in counterbalanced order. Each session included a practice run of 25 trials. A baseline value was then obtained using an FD amplitude designed to produce a response exceeding the previously established threshold by 15 - 25%. After a 5-min pause, the adaptation run was initiated. Here, the pair of stimuli to be discriminated, circle and FD, were preceded, on each of the 50 trials, by an adaptation stimulus whose duration was 500 msec. The interval between adaptation and test stimuli was 2 sec.

Post-adaptation runs were implemented to obtain an estimate of the duration of the adaptation effect. The procedure was identical to pre-adaptation baseline determinations. The first post-adaptation run took place 2 min after the end of adaptation. The second and subsequent post-adaptation runs were conducted at 12, 22, 32 minutes following adaptation, until the baseline value was approximated.

(2) Frequency specificity studies using the staircase method.

A variant of the forced-choice staircase method (Wetherill & Levitt, 1965) was introduced for the following reasons: (1) The data obtained with the method of constant stimuli appeared noisy, despite the reputedly stable results usually obtained from the method of constant stimuli accompanied by the forced-choice procedure. (2) The amount of time consumed by the method of constant stimuli was prohibitive. Each threshold estimate took up to 45 min, as compared to the staircase method which permitted a threshold determination in 5 min. Given that the still more rapid method of adjustment could not be implemented on the 11/23 system, the staircase method was selected instead, and no post-adaptation runs were conducted.

The staircase procedure was implemented as follows: For the baseline run, a tone was delivered, preceding the 2 observation intervals by 250 msec. The standard and comparison stimuli, each 50 ms long, were separated by an interval of 500 ms. The comparison stimulus was presented at random in the first or second stimulus interval. After each pair of stimuli had been presented, the subject indicated which interval had contained the signal and was immediately provided feedback (a double tone) for a correct choice. The amplitude of the comparison stimulus (cloud) was increased



by one-quarter octave after each incorrect response, decreased by the same amount after three consecutive correct responses, and left unchanged otherwise. A block of trials was terminated after 50 trials, or 8 reversals of the amplitude's direction of change (if this had not been accomplished within 50 trials). The first reversal was considered to have occurred after the first series of correct responses following an error. The mean of the eight amplitude values at which reversals had occurred was taken as the estimate of the threshold, which corresponds to the amplitude at which the probability of a correct response is 0.79 (Wetherill & Levitt, 1965).

After an interval of 5 min, an adaptation run was conducted. It was identical to the baseline run except that each discrimination between FD and standard stimulus was preceded by an adaptation figure. This was an FD at one of the following frequencies, 4, 8, or 12. Test frequencies were set at plus or minus one octave relative to the adapting frequency. The amplitude of the adapting frequency was set at 3-4 octaves above the threshold for that frequency. This had proved to be an optimal level for impairing discrimination (elevating baseline values) in earlier studies. After a warning tone, the adapting figure was exposed for 500 msec, with a 2-sec interval preceding the onset of test stimuli. The rest of the procedure was identical to that described for the baseline run.

### 3. Third Experimental Series: Interocular Transfer

Having established the necessary baseline values and optimal adaptation parameters through extensive testing of individual observers, group studies were initiated with naive subjects. These experiments included an assessment of the interocular transfer of adaptation effects, using 7 observers.

The typical procedure for evaluating several properties of adaptation in group studies, such as its duration and its interocular effectiveness was as follows: The forced-choice staircase method was used initially to obtain an estimate of the observer's threshold. A suprathreshold stimulus was then selected which provided a baseline value of 60-75 % correct discriminations over 50 trials. Five min later, an adaptation run was initiated. This consisted of a block of 50 trials in which the test stimuli used in the baseline session were preceded by an adaptation stimulus of optimal amplitude for that frequency. The duration of the adaptation stimulus was again 500 msec, and the interval between adaptation and test stimuli was 2 sec.

The number and timing of post-adaptation runs depended upon the aspect of adaptation being evaluated. In general, however, the first post-adaptation test took place 5 min following the end of the adaptation session. This procedure was employed for testing interocular transfer, except that the baseline values were obtained monocularly, with one eye occluded with a translucent mask, while the adaptation session was the only run conducted with the initially unexposed eye. That is, both baseline and post-adaptation runs were conducted with one eye exposed, while the adaptation run was conducted with the other (usually non-dominant) eye. The FD frequency used for assessing interocular transfer was frequency 8, at a moderately high amplitude.

## RESULTS

### I. First experimental series: Tuning curves.

The curve illustrated in Fig. 34 was obtained from 4 observers who were required to discriminate between a circle and an FD, each at 1.3 degrees of visual angle. These were dark patterns on a lighter background. A relatively systematic relationship was obtained between FD amplitude at discrimination threshold and FD frequency. The curve, however, has a notch in the vicinity of frequencies 6 and 8, which is apparent in the individual functions for observers BM and IA in Figure 35. If the FD discrimination threshold function is presumed to represent the envelope of a set of individual functions or channels, as our frequency specificity results (see below) might suggest, then the anomaly seen at frequencies 6 and 8 may represent a differential sensitivity on the part of some channels (again, see below).

In Figure 35, the stimuli were light patterns on darker ground, also with 1.3 degrees of angular subtense. While observer BM had the higher acuity (Snellen 20/15), his thresholds were higher (for the lower-frequency FDs). This would indicate an absence of relationship, as anticipated, between FD detection and visual acuity.

Changing luminance relations between figure and ground did not affect the results: in Fig. 36, the curve obtained for a contrast value of +.71 involved a light pattern on dark background, while the curve for -.37 represents results from a dark figure on lighter ground.

While the question of size invariance can be approached in several ways, Fig. 37 demonstrates a thorough overlap for curves obtained with patterns of 1.3 and 2.6 degrees of visual angle. The significance of departures from linearity of points for the smaller patterns is as yet unclear. It was felt that before questions of invariance were to be more thoroughly examined, the effects of FD interactions needed to be better understood. Such interactions were evident in studies of thresholds, even when stimulus durations and interstimulus intervals had been lengthened. Studies of adaptation with FDs promised a clearer picture of responsiveness to FD stimuli.

## II. Second Experimental Series: Adaptation and Frequency Specificity

### (1) Time Course and Duration of Adaptation

A reminder of the procedure used for the first series of studies on adaptation will emphasize the remarkable strength of the adapting effects of some FDs. After obtaining a baseline value, a series of 50 trials are conducted where, for each trial, a 500 msec adapting stimulus is followed, 2 sec later, by a discrimination between an FD and a circle. All told, only 25 sec of adaptation are administered. Fig. 38 demonstrates, for two observers, that the peak baseline elevation, for optimal adapting amplitudes, does not occur until 12-16 min later. The intervening test run at 2-6 minutes could not be held responsible for the increased magnitude of the effect 5 min later, or adaptation would have been sustained, rather than decaying after the same interval of time in the next post-adaptation run. Furthermore, more recent studies have shown that delaying the first post-adaptation run often results in a maximal effect at the first run (5-9 min post-adapt.). In effect, an initial post-adaptation test conducted very soon after the adaptation run may hinder, rather than exaggerate, the emergence of the adapting effects. Each point in the curve for BM and IA in Fig. 38 was the average of 3 separate tests for BM and 2 such tests for IA.

It should be noted that it is most unusual to find an aftereffect increasing in strength over time, following adaptation. Although figural aftereffects have been reported to attain a maximum up to 160 msec after the presentation of an I object (Oyama, 1978), it is largely in the domain of contingent aftereffects that longer intervals have been reported. Favreau (1976) tested the color-contingent motion aftereffect both immediately and 7-min after adaptation. She found an increase in the duration of the MAE on the delayed

test, and noted that the color-contingent MAE appeared to be the only kind of negative aftereffect which showed greater strength with time after adaptation.

## (2) Adaptation and Frequency Specificity

Prior to addressing the question of frequency specificity, it was clearly necessary to obtain estimates of the most effective amplitudes for every frequency which we wished to evaluate. This was done for 3 observers whose results indicated that amplitudes at 3 - 4 octaves above threshold for the adapting frequency produced the highest degree of threshold elevation for that frequency.

The first indication that adaptation using FD stimuli was frequency-specific was provided by 2 observers who were tested on frequencies 8 and 6 after having been adapted to frequency 8. Fig. 39 illustrates that, while the predominant effect on test FD 8 was adaptation, or an elevation of baseline values, an effect of opposite sign, facilitation, was seen for test FD 6. This last suggests the operation of disinhibition, and therefore implies the existence of an inhibitory action in FD adaptation. We have pursued this implication in additional studies and will provide the results in a later report (manuscript in preparation).

While the issue of channels of feature analysis may be inherently problematic (Kelly and Burbeck, 1984), Fig. 40 provides evidence for frequency specificity in curvature analysis. Several points must be noted for this figure, obtained by means of the staircase procedure: These curves were generated using the "response" rather than the "sensitivity" method (e.g., Swift & Smith, 1982). They therefore do not necessarily reflect either the exact shape or bandwidth of FD adaptation effects. Nevertheless, they do indicate a differential response to FD frequency. This very likely reflects a selective response to boundary curvature. No claim can be made regarding the existence of independent FD "channels". Nor is it necessary. A great deal of confusion has arisen in the spatial frequency literature as a result of the inference that the visual system processes information through independent spatial frequency channels. There is no evidence here that independent FD channels are responsible for the data observed.

Peak threshold elevation effects (which are all statistically significant) are not equally marked. The degree of adaptation appears to be much greater for frequency 8 for both observers. This could account for the inflection seen in threshold curves (Figs. 34-37). Testing the adapting effect of frequency 6 on one of these observers also produced a high peak effect, tending to confirm this interpretation. If these curves reflect the operation of curvature "channels", it appears that not all of these channels are equally sensitive, a fact more evident here than in the threshold curves.

### III. Third Experimental Series: Interocular Transfer following Adaptation with FDs.

Using frequency 8 at its optimal adaptation amplitude, 7 observers demonstrated interocular transfer of the adaptation effect, as illustrated by mean values for the group in Fig. 41. The first (solid) bar reflects the mean percent correct responses (corrected for chance) at baseline (67 %) for the exposed eye. An adaptation run implemented for the contralateral eye is not illustrated. A series of post-adaptation runs was conducted at 5, 15, and 25 min following the end of adaptation. The mean percentages correct for the group were 47%, 43%, and 62% for these post-adaptation runs. The first and second post-adapt values differed significantly from baseline (correlated t-tests,  $p < .05$  and  $p < .01$ , respectively). The maximal baseline elevation was obtained at 15 min after the end of adaptation, indicating a long time course for interocular effects. The baseline value was restored by 25 min following adaptation.

This demonstration of interocular transfer of FD adaptation, as well as its duration, strongly suggests that the locus of FD adaptation is beyond the retina.

### SUMMARY

The possibility that boundary curvature is analyzed by brain regions long associated with shape perception has been raised by recent physiological studies in IT cortex (Schwartz et al, 1983). The existence of curvature detectors has also been suggested by the psychophysical experiments of, for example, Riggs (1973, 1974). The controversy surrounding the interpretation of Riggs's findings largely centers on the role of orientation detectors, whether these are necessary and sufficient for the analysis of curvature

(e.g., MacKay and MacKay, 1974; Sigel and Nachmias, 1975). Schwartz (1980a) has suggested a possible mechanism whereby local edge information in striate and prestriate cortex might be filtered to produce a global analysis of boundary curvature. This formulation bypasses the controversy mentioned above and suggests the plausibility of adopting an approach using FD stimuli for the psychophysics of boundary curvature.

The most prominent features of our studies with FDs are:

(1) the duration of the effects of relatively brief adapting FDs.

(2) the interocular transfer of FD adaptation.

(3) the frequency specificity of the effects of FD adaptation.

These were the major effects we had set out to demonstrate in our proposal. They strongly suggest the utility of the FD approach for the quantitative analysis of boundary curvature in the human observer.

## BIBLIOGRAPHY

- Ahlfors, L., Complex Analysis, McGraw\_Hill, New York, 1966.
- Albright, T., R. Desimone, and C. G. Gross, "Columnar organization of directionally selective cells in visual area MT of the macaque monkey," J. Neurophysiology, 1984.
- Alter, I. and E.L. Schwartz, "Discrimination thresholds and adaptation to periodic boundary curvature (Fourier Descriptor) stimuli.," Inv. Ophthalmol. & Vis. Sci. (Supp.), vol. 25, p. 200, 1984. (Abstract)
- Attneave, F., "Some informational aspects of visual perception.," Psychological Review, vol. 61, pp. 183-193, 1954.
- Barlow, H.B., D.I.A. MacLeod, and A. van Meeteren, "Adaptation to gratings: No compensatory advantages found.," Vision Res., vol. 16, pp. 1043-1045, 1976.
- Barlow, H. B., "Single units and sensation: a neuron doctrine for perceptual psychology," Perception, vol. 1, pp. 371-380, 1972.
- Blakemore, C. and R. Over, "Curvature detectors in human vision?," Perception, vol. 3, pp. 3-7, 1974.
- Blakemore, C., R.H.S. Carpenter, and M.A. Georgeson, "Lateral inhibition between orientation detectors in the human visual system.," Nature, vol. 228, pp. 37-39, 1970.
- Braccini, C., G. Gambardella, G. Sandini, and V. Tagliasco, "A model of the early stages of the human visual system: functional and topological transformations in the peripheral visual field," Biological Cybernetics, vol. 44, pp. 47-58, 1982.
- Brocker, T., Differential Germs and Catastrophes, Cambridge University Press, Cambridge, 1975.
- Brousil, J. K. and D. R. Smith, "A threshold logic network for shape invariance," I.E.E.E. Transactions on Computers, vol. EC-16, pp. 818-828, 1967.
- Burbeck, C.A. and D. Regan, "Independence of orientation and size in spatial discriminations.," J. Opt. Soc. Am., vol. 73, pp. 1691-1694, 1983.

- Burt, Peter J. and Edward H. Adelson, "The Laplacian Pyramid as a Com[act Image Code," IEEE Transactions Communcations, vol. COM-31, pp. 532-536, 1983.
- Campbell, F. W. and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiology, vol. 197, pp. 551-556, 1968.
- Casasent, D. and D. Psaltis, "Position, rotation and scale invarianct optical correlation," Applied Optics, vol. 15, pp. 1793-1799, 1976.
- Cavanagh, P., "Size, rotation and position invariance in the visual system," Perception, vol. 7, pp. 167-177, 1978.
- Cavill, J. and J.O. Robinson, "A color aftereffect contingent on complex pattern features.," Perception & Psychophysics, vol. 19, pp. 454-459, 1976.
- Daniel, D. M. and D. Whitteridge, "The representation of the visual field on the cerebral cortex," J. Physiology, vol. 159, pp. 203-221, 1961.
- Desimone, R., E. L. Schwartz, T. D. Albright, and C. G. Gross, "Inferior temporal neurons selective for shape," Investigative Ophthalmology and Visual Science ARVO Supplement, 1982.
- DeValois, R., D. Albrecht, and L. G. Thorell, "Spatial tuning of LGN and cortical cells in monkey visual system," in Spatial Contrast, ed. H.S.Spekreisje and L.H. van der Tweel , pp. 60-63, North Holland, Amsterdam, 1977.
- deValois, R. and K. deValois, "Spatial vision," Annual Rev. Psych., vol. 31, pp. 309-341, 1980.
- Dodson, V. G., "Neuronal circuits capable of generating visual cortex," Perception, vol. 9, pp. 411-434, 1980.
- Dow, B. M., A. Z. Snyder, R. G. Vautin, and R. Bauer, "Magnification factor and receptive field size in foveal striate cortex of monkey," Exp. Brain Res., vol. 44, pp. 213-228, 1981.
- Drasdo, N., "The neural representation of visual space," Nature, vol. 256, pp. 554-556, 1977.
- Eggermont, P. P. B., "Logarithmic spiral grids, computer assisted tomography and abel type integral equations," pre-print,



1981.

- Ermentrout, G. B. and J. D. Cowan, "A mathematical theory of visual hallucinations patterns," Biological Cybernetics, vol. 34, pp. 137-150, 1977.
- Favreau, O.E., "Interference in colour-contingent motion after-effects.," Quart. J. exp. Psychol., vol. 28, pp. 553-560, 1976.
- Fu, K. S., Syntactic Pattern Recognition, Academic Press, New York, 1974.
- G.T.Herman,, Image Reconstruction from Projections, Academic Press, New York, 1980.
- Gataas, R., C. G. Gross, and J. Sandell, "Visual topography of V2 in the macaque," J. Comp. Neurology, vol. 188, pp. 301-315, 1981.
- Gibson, J. J., The Senses Considered as Perceptual Systems, Houghton-Mifflin, Boston, 1966.
- Goldman, P. and W. Nauta, "Columnar distributions of cortico-cortical fibers in frontal, association, limbic and motor cortex of developing rhesus monkey," Brain Res., vol. 122, pp. 393-413, 1977.
- Gross, C. G., C. E. Rocha-Mirand, and D. Bender, "Visual properties of neurons in infr," J. Neurophysiology, vol. 235, pp. 96-111, 1972.
- Hartline, H. K. and F. Ratliff, Studies On Excitation and Inhibition in the Retina, Rockefeller University Press, New York, 1974.
- Hayman, C.A.G. and L.G. Allan, "A reevaluation of angle-contingent color aftereffects.," Perception & Psychophysics, vol. 28, pp. 61-67, 1980.
- Held, R., "Adaptation to rearrangement and visual-spatial aftereffects.," Psychologische Beitrage, vol. 6, pp. 439-450, 1962.
- Helmholtz, H. von, Treatise On Physiological Optics, Optical Soc. of America (translated by J.P.C. Southall, New York, 1925.
- Hubel, D. H. and J. Horton, Abstract-Society for Neuroscience, 10th Annual Meeting, 1980.

- Hubel, D. H., "Brain mechanisms of vision," Scientific American, vol. 241, pp. 150-163, 1979.
- Hubel, D. H. and D. C. Freeman, "Projection into the visual field of ocular dominance columns in macaque monkey," Brain Res., vol. 122, pp. 336-343, 1977.
- Hubel, D. H. and T. N. Weisel, "Receptive fields, binocular interaction, and functional architecture in cats visual cortex," J. Physiol., vol. 160, pp. 106-154, 1962.
- Hubel, D. H. and T. N. Weisel, "Sequence regularity and geometry of orientation columns in the monkey striate cortex," J. Comp. Neurol., vol. 158, pp. 267-293, 1974.
- Hubel, D. H. and T. N. Weisel, "Stereoscopic vision in macaque monkey," Nature, vol. 225, pp. 41-44, 1970.
- Hummel, R., "Zero crossing and the heat equation," Tech Report, Courant Institute of Math. Sciences, 1984.
- III, C.F. Stromeyer, "Curvature detectors in human vision?," Science, vol. 184, pp. 1199-1200, 1974.
- Jain, H. K., "Optical Flow and logarithmic mapping," Proceedings of Toronto Conf. Motion Perception, 1983.
- John, E.R. and Eric L. Schwartz, "The neurophysiology of information processing and cognition," Annual Review of Psychology, vol. 29, pp. 1-29, 1978.
- Julesz, B., Foundations of Cyclopean Perception, Univeristy of Chicago Press, Chicago, 1971.
- Kelly, D.H. and C.A. Burbeck, "Critical problems in spatial vision.," CRC Critical Reviews in Biomedical Engineering, vol. 10, pp. 125-177, 1984.
- Kelly, D.H., "Pattern detection and the two-dimensional Fourier transformation: Flickering checkerboards and chromatic mechanisms.," Vision Res., vol. 16, pp. 277-289, 1976.
- Koenderink, J. J. and A. J. Van Doorn, "Local structure of movement parallax in the plane," J. Opt. Soc. America, vol. 66, pp. 717-723, 1976.

- Koenderink, J. J. and A. J. van Doorn, "Visual detection of spatial contrast: influence of location in the visual field, target extent and illuminance level," Biological Cybernetics, vol. 30, pp. 157-167, 1978.
- Kovacsny, J. and R. Josephs, "Symmetric edge operators," Proc. I. R. E., vol. 244, 1958.
- Kurtenbach, W. and S. Magnussen, "Inhibition, disinhibition, and summation among orientation detectors in human vision.," Exp. Brain Res., vol. 43, pp. 193-198, 1981.
- Lettvin, J., H. Maturana, W. Pitts, and W. McCulloch, "What the frogs eye tells the frogs brain," Proc. I.R.E., vol. 47, pp. 51-55, 1959.
- LeVay, S., D. H. Hubel, and T. N. Wiesel, "The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain," J. Comp. Neurol., vol. 159, pp. 559-576, 1975.
- Livingstone, Margaret S. and David H. Hubel, "Anatomy and physiology of a color system in the primate visual cortex," Journal of Neuroscience, vol. 4, pp. 309-356, 1984.
- Longuet-Higgins, H. C. and K. Prazdny, "The interpretation of a moving retinal image," Proc. Roy. Soc. B, vol. 208, pp. 385 - 397, 1980.
- MacKay, D. M., "Clues to the site of origin of the complementary image," Nature, vol. 279, p. 553, 1979.
- MacKay, D.M. and V. MacKay, "Do curvature-contingent chromatic aftereffects require "detectors for curvature"?, " Vision Res., vol. 14, pp. 1285-1287, 1974.
- MacKay, D.M., "Moving visual images produced by regular stationary patterns.," Nature, vol. 180, pp. 849-850, 1957.
- MacKay, D.M., "Perception and brain function.," in The Neurosciences: Second Study Program, ed. F.O. Schmitt, pp. 303-316, The Rockefeller University Press, New York, 1970.
- Marr, D., Proceedings of the Austing Feature Extraction Conference, Academic Press, New York, 1979.
- Marr, D. and T. Poggio, "A computational theory of human stereo vision," Proc. Roy. Soc. B, vol. B277, pp. 301-328, 1979.
- Massone, L., G. Sandini, and V. Tagliasco, "A form invariant topological mapping strategy for 2D Shape Recognition," Preprint, 1983.

- Meyerson, J., P. B. Manis, F. Mieszin, and J. M. Allman, "Magnification factor in striate cortex and retinal ganglion cell layer for owl monkey: A quantitative comparison," Science, vol. 98, pp. 855-857, 1977.
- Mishkin, M. A., "A sequential neural model of visual perception, memory and affect," in Karl Lashley Lectures: Neuropsychology After Lashley Queens College, New York, 1979.
- Noton, L. and L. Stark, "Visual scan paths," Sci. American, vol. 178, 1976.
- Oyama, T., "Figural aftereffects," in Handbook of Sensory Physiology, ed. H.L. Teuber, vol. VIII, pp. 570-592, Springer-Verlag, New York, 1978.
- Pennington, K. S., P. M. Will, and G. L. Shelton, "Grid coding: a technique for extraction of differences from scenes," Optics Communications, vol. 2, pp. 113-119, 1970.
- Persoon, E. and K. S. Fu, "Shape description using Fourier descriptors," I.E.E.E. Trans. on Systems, Man and Cybernetics, vol. SMC-7, pp. 170-179, 1979.
- Poggio, G. and B. Fischer, "Binocular interaction and depth sensitivity in striate cortex and pre-striate cortex of behaving rhesus monkey," J. Neurophys., vol. 40, pp. 1392-1405, 1977.
- Pollen, D., J. R. Lee, and J. H. Taylor, "How does the striate cortex begin the reconstruction of the world," Science, vol. 173, pp. 74-77, 1971.
- Pratt, W. K., Digital Image Processing, Wiley, New York, 1978.
- Ransom-Hogg, A. and L. Spillman, "Receptive field size in fovea and periphery of the light and dark adapted retina," Vis. Res., vol. 20, pp. 221-228, 1980.
- Regan, D., K. Beverly, and M. Cynader, "The visual perception of motion in depth," Sci. AM., vol. 241, pp. 122-133, 1979.
- Richards, W., "Fortification illusions associated with migraine headache," Scientific American, vol. 78, pp. 115-120, 1971.
- Riggs, L.A., "Curvature as a feature of pattern vision," Science, vol. 181, pp. 1070-1072, 1973.
- Riggs, L.A., "Curvature detectors in human vision?," Science, vol. 184, pp. 1200-1201, 1974.

- Rovamo, J. and V. Virsu, "An estimation and application of human cortical magnification factor," Exp. Brain Research, vol. 37, pp. 495-510, 1979.
- Sakitt, B., "Why the cortical magnification factor in rhesus monkey cannot be isotropic," Vis.Res., vol. 22, pp. 417-421, 1982.
- Schwartz, E. L., "Anatomical and physiological correlates of visual computation from striate to inferotemporal cortex," IEEE Transactions on Systems, Man and Cybernetics, vol. SMC 14, pp. 257-271, 1984.
- Schwartz, E.L., "Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding.," Vision Res., vol. 20, pp. 645-669, 1980.
- Schwartz, E.L., R. Desimone, T.D. Albright, and C.G. Gross, "Shape recognition and inferior temporal neurons.," Proc. Natl. Acad. Sci., vol. 80, pp. 5776-5778, 1983.
- Schwartz, Eric L., "Afferent geometry in the primate visual cortex and the generation of neuronal trigger features," Biological Cybernetics, vol. 28, pp. 1-24, 1977.
- Schwartz, Eric L., "Columnar architecture and computational anatomy in primate visual cortex: segmentation and feature extraction via spatial frequency coded difference mapping.," Biological Cybernetics, vol. 42, pp. 157-168, 1982.
- Schwartz, Eric L., "Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding," Vision Research, vol. 20, pp. 645-670, 1980.
- Schwartz, Eric L., "Cortical anatomy and size invariance," Vision Research, vol. 18, pp. 24-58, 1983.
- Schwartz, Eric L., "Cortical mapping and computational anatomy: a projective invariant recursive flow model of visual coding.," Advances in Physiological Sciences. Edited by E. Grastyan and P. Molnar, Pergamon Press, New York, 1980.
- Schwartz, Eric L., "The development of specific visual projections in the monkey and the goldfish: outline of a geometric theory of receptotopic structure.," Journal of Theoretical Biology, vol. 69, pp. 655-685, 1977.
- Schwartz, Eric L., David R. Christman, and Alfred P. Wolf, "Human primary visual cortex topography imaged via positron emission tomography," Brain Research, vol. In press, 1983.

- Schwartz, Eric L., "On the mathematical structure of striate cortex topography," Science, vol. In press, 1984.
- Schwartz, Eric L., "A quantitative model of the functional architecture of human striate cortex with application to visual illusion and cortical texture analysis," Biological Cybernetics, vol. 37, pp. 63-76, 1980.
- Schwartz, Eric L., "Quantitative models of functional architecture in striate cortex," Models of Striate Cortex, ed. by R. Rose and D. Halton, London, England, 1984.
- Schwartz, Eric L., Robert Desimone, Tom Albright, and Charles G. Gross, "Shape recognition and inferior temporal neurons," Proceedings of the National Academy of Sciences, 1983.
- Schwartz, Eric L., "Spatial frequency analysis, cortical anatomy and size invariance," Perception, vol. 10, pp. 455-468, 1981.
- Schwartz, Eric L., "Spatial mapping and spatial vision in monkey and human visual cortex," In: Sensation, Perception and Cognition: A Festschrift for Ivo Kohler, ed. by L. Spillman and B. Wooten, Erlbaum Press, New York, 1983.
- Schwartz, Eric L., "Spatial mapping in the primate sensory projection and relevance to perception," Biological Cybernetics, vol. 25, pp. 181-195, 1977.
- Schwartz, Eric L., "Spatial mapping in the visual system," Journal of the Optical Society of America, vol. 68, p. 1371, 1978.
- Segal, L. A., Mathematics Applied to Continuum Mechanics, MacMillan, New York, 1977.
- Shenker, P. A., "Toward the robot eye: In search of an isomorphic image transform for machine vision," Tech. Report Brown University ENG RRV 81-1, 1981.
- Sigel, C. and J. Nachmias, "A re-evaluation of curvature-specific chromatic aftereffects," Vision Res., vol. 15, pp. 829-836, 1975.
- Skowbo, D., B.N. Timney, T.A. Gentry, and R.B. Morant, "McCollough effects: Experimental findings and theoretical accounts," Psychol. Bull., vol. 82, pp. 497-510, 1975.
- Stensaas, S. S., D. K. Eddington, and W. Dobelle, "Human striate cortex," J. Neurosurgery, vol. 40, pp. 747-758, 1974.
- Stevens, R. J., A. F. Lehar, and F. H. Preston, "Manipulation and presentation of multidimensional data using the Peano scan," IEEE PAMI, pp. 520-532, 1983.

- Swift, D.J. and R.A. Smith, "An action spectrum for spatial-frequency adaptation," Vision Res., vol. 22, pp. 235-246, 1982.
- Teuber, H.-L., "Perception," in Handbook of Physiology. Section 1: Neurophysiology, vol. III, pp. 1595-1668, American Physiological Society, Washington, 1960.
- Tootell, R. B., M. Silverman, and R. deValois, "Deoxyglucose analysis of retinotopic organization in primate striate cortex," Science, vol. 218, pp. 902-904, 1982.
- Tootell, R. B., M. S. Silverman, and R. L. DeValois, "Spatial frequency columns in primary visual cortex," Science, vol. 214, pp. 813-815, 1981.
- Tusa, R. J., L. A. Palmer, and A. C. Rosenquist, "The retinotopic organization of area 17 (striate cortex) in the cat," J. Comp. Neurol., vol. 177, pp. 213-236, 1978.
- Tyler, C. W., "A new class of visually induced phosphenes," Vis. Res., vol. 15, pp. 403-420, 1979.
- Van Essen, D. C., W. T. Newsome, and J. H. R. Maunsell, "The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability," Vis. Res., vol. 24, pp. 429-448, 1984.
- Virsu, V. and J. Rovamo, "Visual resolution, contrast sensitivity and the cortical magnification factor," Exp. Brain Res., vol. 37, pp. 474-494, 1979.
- Weiman, C. F. and G. Chaiken, "Logarithmic spiral grids for image processing and display," Comp. Graphics and Image Processing, vol. 11, pp. 197-226, 1979.
- Westheimer, G., "The spatial grain of the perifoveal visual field," Vis. Res., vol. 22, pp. 157-162, 1981.
- Wetherill, G.B. and H. Levitt, "Sequential estimation of points on a psychometric function," The Brit. J. math. stat. Psychol., vol. 18, pp. 1-10, 1965.
- White, K.D. and L.A. Riggs, "Angle-contingent color aftereffects," Vision Res., vol. 14, pp. 1147-1154, 1974.
- Winston, P., The Psychology of Computer Vision, M.I.T. Press, Cambridge, Mass., 1975.
- Zahn, C. R. and R. Z. Roskies, "Fourier descriptors for plan closed curves," I.E.E.E. Trans. on Computers, vol. C-21, pp. 269-281, 1972.
- Zucker, S. W. and R. Hummel, "An application of relaxation labeling to line and curve enhancement," PAMI, 1979.

## Figure 1.

This figure shows a synthetic sphere, with two light sources, smooth shading, and antia-aliasing, constructed via our solids modeling software.

## Figure 2A

This is a reconstruction, via the same software as figure 1, of the surface of striate cortex of a macaque monkey. The deep sulcus to the right is the lateral calcarine sulcus. The fovea is to the extreme left, and the vertical meridian forms the upper and lower borders of the figure. This is the "opercular" surface, representing approximately the region 0-9 degrees of visual field.

## Figure 2B.

This is a wire frame model of the brain surface of figure 2A.

## Figure 3A,B

This figure is a photograph from an HO scale model (1:86) which was used for many of the scenes in this report. These two figures are a stereo pair: They were shot with a reticle in the camera focused on the center of the star of the tank, but from two slightly different positions. Thus, there is binocular disparity in the scene. This can be verified by fusing the double image as with a Julesz random dot stereogram. The bus was in front of the fixation plane, and the small car was behind the fixation plane by an equal amount.

## Figure 4.

This is similar to figure 3, except that motion was introduced. The tank was moved between two frames. Then, the two slightly different images of the tanks were interlaced. Where the mis-match occurs, there is a prominent "rippling" of the contours of the tank. This is a simple and well known cue for motion Pennington, which is similar to cues introduced by the interlacing of left and right eye images at the level of striate cortex ( see figures 17,18, and 19 below).

## Figure 5

This figure illustrates display of full color frames using only 8 bits of image memory, and lookup tables. The red, green and blue components of a movie scene are shown; the lower right of the four frames is the composite full-color image. The quality of the color is not terribly high, due to the several generations of reproductions here. On our monitor screen, we felt that the reproduced color was very close to that of the slide. The red, green and blue components were digitized from a color slide using a color enlarger and an 8 bit CCD camera.



Below is an example of a Peano ( or Hilbert) curve. This curve snakes through 3-space, touching every point one and only once. It has been used, as described in the text, to rebin our 3x 5bit color space to a single 8 bit image, with lookup table values which provide accurate color.

Figure 6

This figure illustrates the map function  $\log(z+0.3)$ , and its current agreement with several different sources of data. The top figure shows a "retinal" semi-circle of equal spaced rings at one degree intervals, and on the right shows the map of this figure under the map function  $\log(z+0.3)$ . Underneath is shown this map superimposed over the operculum reconstruction of striate cortex of Levay 1975. In the center is shown the fit of Dow et al Dow 1981 to striate cortex magnification. Superimposed is the calculation from the map function  $\log(z+0.3)$ .

The top two figures were part of a previous publication, and were also submitted in our previous AFOSR proposal. The bottom figure was constructed about six months later, when Tootel 1982 published their 2-deoxyglucose mapping of striate cortex. Shown on the left is the map function  $\log(z+0.3)$ , calculated for the rings and rays used by Tootel et al. In the center bottom is shown the Tootel map, and on the right is a superposition of the theoretical model  $\log(z+0.3)$  and Tootels data.

Curiously, Tootel et al claimed in their paper that there were significant disagreements between their data and the complex log model, although they apparently did not bother to superimpose the data and the model. They find discrepancies in the lower right corner of the map. But this is also the area where the curvature of the cortex is greatest, and where there is most likely to be distortion due to their fairly crude "squashing" of the cortex between glass slides. In fact, the lower edge of the vertical meridian representation is lost in their data( the area which is marked by square boxes) due to this curvature.

In summary, it seems, based on the three independent pieces of data shown above, that the map function  $\log(z+0.3)$  may be a useful approximation to cortical topographic map structure, which is consistent with most aspects of current data, and which allows a convenient modeling of striate cortex topography.

Figure 7.

A photo of the space shuttle Columbia. Nine points are marked by crosses; these are the fixation points to be

used in later discussion. The small squares are the maps obtained from the function  $\log(z+0.3)$ , centered at each of the fixation points indicated in the figure. The scale of the log images to the original scene is correct. The tiny size of the cortical images is due to the tremendous data compression due to the log mapping. The next figure shows an expansion of one of the log images, which are too small to be seen in this figure.

Figure 8

A single log mapped view along the space shuttle, centered between the "d" and "S" of "United States". This view is obtained by taking the cortical map, and backprojecting it to the "retina" with the inverse map function  $\text{cexp}(w) - 0.3$ .

The blocky quality of the image is due to our current log mapping methods, which do not perform bilinear interpolation on the images. We are currently implimenting this "cosmetic" improvement, but will probably continue to work with the current integer software, since it is very fast, and gives an accurate representation of the map structures.

Figure 9.

An image from "HO" scale, of a toy tank. The log mapped image is at the top, using the map function  $\log(z+1.7)$ . The value 1.7 is favored by most modelers of cortical topography ( Hubel and Weisel, Drasdo and Virsu and Rovamo). We don't feel that it is correct, but it is easier to work with, due to the much smaller expansion at the center of the field.

Figure 10.

Same as above, but closer to the tank.

Figure 11

Same as above, but still closer to the tank.

Figure 12.

One of the tank images from above, in "retinal" view, for the map function  $\log(z+1.7)$ . Note the relatively mild loss of detail with eccentricity.

Figure 13

Same as above, but with map function  $\log(z+0.3)$ . This is a much stronger change in magnification, and probably closer to the true value for monkeys, and perhaps for humans. Due to the 512x512 resolution of our current image software, we cannot obtain the true detail in the center (region of the star) that should be there.

It is interesting to step back five or six feet from this image. The details of the tanks, its wheels, etc. may then be clearly seen. This space variant fall of in resolution is similar to the famous figure of Lincolns face of Harmon and Julesz, which is a constant resolution figure which gradually comes into detail as one steps back, or defocuses ones eyes.

Figure 14.

This figure shows a Gaussian pyramid, after Burt and Adelson 1983 The space shuttle image is blurred by convolution with a gaussian kernel, then sub-sampled, then blurred again, then sub-sampled, etc.

Figure 15.

This figure shows a Laplacian Pyramid. This is formed by subtracting adjacent levels of the Gaussian Pyramid above. This gives a difference of gaussian edge enhancer. Hummel Hummel 1984 has recently shown that this is a discrete approximation to convolution with a Laplacian, a fact which was well known on less formal grounds by most workers in vision. The Laplacian Pyramid provides a convenient summary of the multi-resolution and circularly symmetric lateral inhibition of retinal spatial processing.

This figure shows an example of ocular dominance columns obtained in collaboration with Dr. B. Merker (NYU). The monkey had one eye removed for one month prior to sacrifice. The interlaced patches of the remaining eye columns are clearly visible in this cytochrome oxidase stain which has been pseduo-colored.

Figure 17.

In this figure, Levays Levay 1975 reconstruction of the macaque ocular dominance column pattern is shown on the left, for the central two degrees of field ( under the map function  $\log(z+0.3)$ )

An image ( 2 degree patch of the space shuttle in the region of "United States") is "anded" with the ocular dominance column pattern to project out image columns. Note, that without a good map function (i.e. one which more or less fits the LeVay data, as well as Tootels data, as in figure 6 above) it would be impossible to construct this image.

Figure 18.

Figure 18 is projected back to the retina using the complex exponential.

Figure 19.

Two images of the space shuttle are interlaced with two different amounts of binocular disparity ( in this

case, merely lateral shift of the "left" and "right" eye image). The disparity on top is 3 pixels, and on the bottom is 5 pixels. The resultant "cortical" simulation was then high pass filtered with a Sobel operator (vertical and horizontal). The resultant "rippling" of the contours of the shuttle is due to mismatch of the contours at the column borders, due to the "disparity". The magnitude of this "ripple" is proportional to the disparity, and is therefore a rough "textural" cue for stereo disparity. On the right is the unfiltered cortical image

Figure 20.

This figure shows the several steps in simulating a "hypercolumn" with a small scale complex log mapping. On the left are shown four grid patterns. Immediately to the right of them is the complex log mapping of the grid. In this mapping, the second of two choices for orienting the two "hemi-spheres" discussed in earlier work is used. Schwartz 1977. Afferent Note that this is meant to simulate a very small patch of cortex, on the order of 1 square mm. Also note that the grid patch on the left is only a part of the patch of grids which were log mapped. In the third column is the log mapped parallel grid lines, with a vertical edge operator, to simulate the effect of directional lateral inhibition, as specified in our original model. This "lateral inhibition" operator was a one directional Sobel operator. In the fourth column is a smoothing of the differentiated image, via a median filter (3x3). Finally, this image is thresholded, at the far right, to simulate neuronal firing. The result is a series of band-like regions, which are interrupted in the center. These regions would represent the areas of maximal response to a pattern of periodic lines, at a given orientation, as on the far left. Thus, both of major aspects of the hypercolumn model of Hubel and Wiesel are displayed: bands of orientation tuned neurons such that the band shifts in one direction, locally, while the stimulating pattern rotates, and a failure of this model in the center of the hypercolumn, which is due to the collapsing of all orientations in that region.

Figure 21

A pseudo colored version of the hypercolumn model illustrated in figure 20. On top, the shifting of the band of orientation tuning, with the rotation of the stimulating pattern, is evident. On the bottom, a similar demonstration is made for orthogonal contours. This property, that orthogonal contours are spatially complimentary at the level of striate cortex, is the basis of a possible explanation of the MacKay complimentary aftereffect. Schwartz 1980. Vision Research which will be simulated with this software in the near

future.

Figure 22A. An array of simulated hypercolumns, produced as in the previous two figures. Beneath is shown data from Livingstone and Hubel 1984 which indicates the spot like pattern of cytochrome oxidase spots in striate cortex, which are the central regions of each hypercolumn which tend to lack strong orientation tuning.

Figure 23

This figure shows a reconstructed view of the space shuttle, which was constructed by "back projecting" 9 individual log mappings, for given fixation points as shown in figure 7 above. This view represents a kind of idealized "minds eye" reconstruction, which presupposes that all spatial information acquired over a period of 9 fixations ( perhaps 1.8 seconds) is retained and "blended" together. In any case, this image represents the maximum information which is taken in over this given set of fixation directions. Note that the detail should be much higher in this figure: the area of the text "United States" on the side of the shuttle was scanned with four fixation points (see figure 7). However, since our 512x512 resolution is not adequate to give full justice to the magnification provided by the map function used ( $\log(z+0.3)$ ), we need to wait for our higher resolution software to be ready to study the full import of this model of scanning.

One final point is that this image was constructed from a DOG (difference of Gaussian filter), applied to each of the cortical images. We have experimented with several different blending algorithms: 1.) direct algebraic averaging of multiple views, 2.) averaging on separate levels of a Laplacian Pyramid, then back projecting, 3.) backprojecting of a high pass filtered log image, as in the present case. Although there are not great differences between the methods, we feel there may be a slight advantage to Pyramid based blending, since images treated in this way have slightly better high frequency detail.

Figure 24.

On top is an image of a town, using three different mapping functions:

- 1.)  $\log(z+1.7)$  for 9 degrees of hemi-field
- 2.) town ("HO" gauge) for 25 degrees of hemifield.
- 3.) town for map function of  $\log(z+0.3)$  and 9 degrees of field

Beneath is the tank image, for the same three map functions.

## Figure 25.

The same as above, but for a closer view point. The motion is toward the star on the tank, which is also the optical axis. A similar relation of "impact point" and optical axis holds for the town scene above. For this simple relation between optical axis and motion axis, there is a simple optical flow condition on the images. This is most clearly seen in figure 31, for a simpler image, but it is also apparent after close inspection of the current figure.

## Figure 26.

The same as above, but for a closer viewing distance/

## Figure 27.

This figure illustrates moving towards one "impact point" and looking at another. In this case, the impact point is the center of the star on the tank. The viewing point is shown by the six crosses on the image.

## Figure 28.

Optical flow on tank image, for condition of moving towards star of tank, but looking at the six different fixation points, between zero and five degrees away from this "impact point".

Figure 29. Same as above, but for further distance from tank.

## Figure 30

Same as above, but for still further distance from tank.

## Figure 31.

A wrench was approached with fixation point along the line of sight. For this simple case of optical flow, and simple image, the size invariant nature of optical flow in the cortical ( complex logarithmic) plane is shown.

## Figure 32.

A left and right eye stereo pair, and the right eye scene with some relative motion, were processed with the Peano scan algorithm.

the stereo pair had a fixation point on the star on the tank. The car had crossed disparity, and the bus had uncrossed disparity of the same magnitude. The stereo pair were fed to the red and green channels of the image. The third frame consisted of the same as the right eye scene, but with the tank slightly moved. This third ( motion) scene was input along the blue channel of the imaging system. Although these images are unprocessed, and not very appropriate tests of our multi-modal approach to segmentation, the difference

contours of the tank were rippled with blue, while its rear part was rippled with yellow. Similarly, the car (and other larger disparity objects in the background) were rippled with green and yellow.

Figure 33

Illustrations of single FDs: beginning from the bottom left, counterclockwise, frequencies 2, 4, 8, 16, 32, and 64 (lobes barely visible), all at amplitude 0.8 in the units of Zahn and Roskies (1972). The FDs used in our studies were outlines, rather than solid figures.

Figure 34

The mean amplitudes and their standard deviations at discrimination threshold for 4 observers are plotted against FD frequency, on logarithmic coordinates. Amplitude units are those of Zahn and Roskies (1972).

Figure 35

FD amplitudes at discrimination threshold are plotted for 2 observers, BM and IA.

Figure 36

Two threshold curves at different contrast ratios are shown for the same observer.  $+0.71$  represents a light pattern on a dark background, while  $-0.37$  refers to a dark pattern on lighter ground.

Figure 37

Two threshold curves at different visual angles, 2.6 and 1.3 degrees, were obtained from the same observer.

Figure 38

Two curves illustrating the course of adaptation following an adapting pattern at frequency 8, amplitude 0.4, were obtained from 2 observers, BM and IA. In both cases, peak baseline elevations are seen 12-16 minutes after exposure to a series of 50 500-msec adapting stimuli

(AD on the abscissa).

Figure 39

The percentage change from baseline for test frequencies 8 and 6 following adaptation at frequency 8 is presented for observer IA. Light bars represent baseline elevation, while dark bars indicate baseline diminution or facilitation. The maximal values for both baseline elevation and facilitation are approximately 60%. While the horizontal axis represents adapting amplitude for frequency 8, the oblique axes reflect time elapsed from the beginning of adaptation. The first bar along each oblique axis is the adaptation

session, where 50 500-msec adapting stimuli are presented concurrently with 50 discrimination trials. The second bar is the first post-adaptation run of 50 discrimination trials extending from 2-6 min following adaptation. Subsequent bars, 3 and 4, are identical post-adaptation trials taking place at 12-16 and 22-26 min following the end of adaptation. Baseline elevations predominate for adaptation and testing at the same frequency (8), while facilitation occurs more often at frequency 6 (except for the adaptation trial itself, particularly at the higher amplitudes).

Figure 40

Frequency specificity is here demonstrated for 2 observers, ALE and CG, following exposure to FD adapting patterns at 3 frequencies:

4 (diamonds), 8 (circles), 12 (triangles). Test frequencies are represented logarithmically along the abscissa.

Fig. 41

Interocular transfer of the effects of adaptation is here shown for a group of 7 observers. The first (solid) bar represents the monocular baseline (mean percent correct discriminations). The adaptation run conducted with the other eye is not illustrated. Post-adaptation mean values at 5, 15, and 25 min following the end of adaptation are shown for the initially exposed (unadapted) eye.

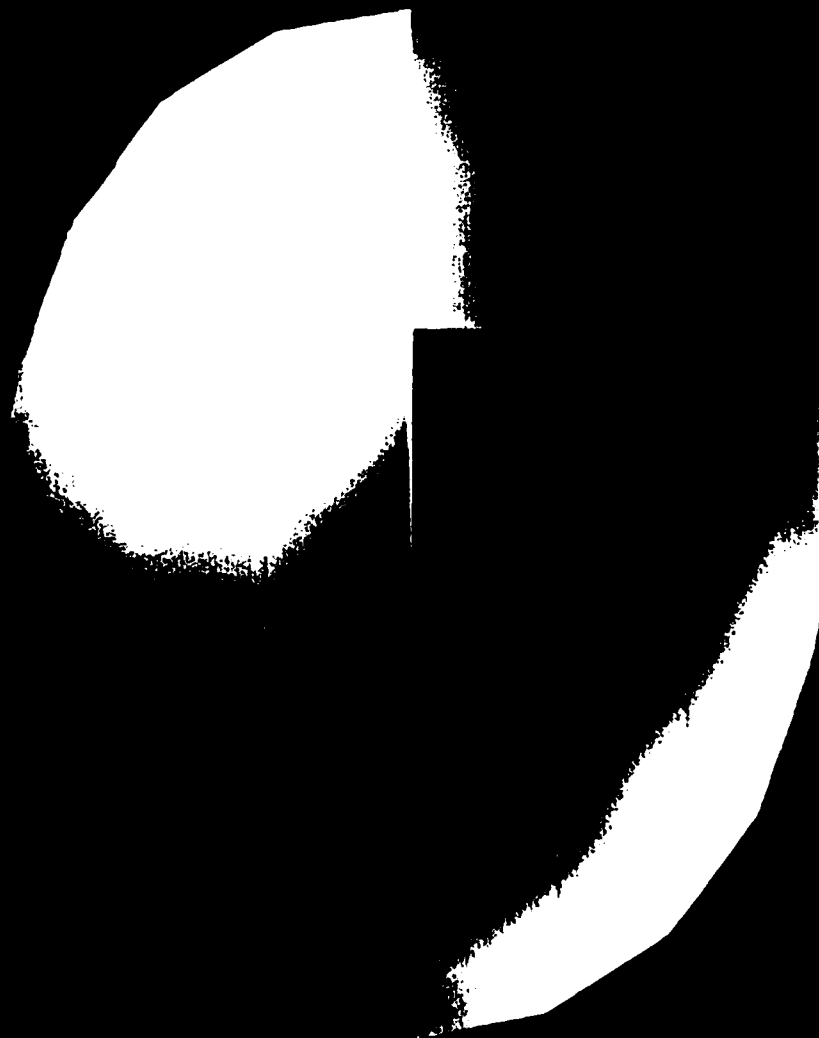


**NOTE:**

ONLY ONE ORIGINAL(AND FIFTEEN COPY SETS)OF FIGURES ARE INCLUDED IN THIS FINAL REPORT. SINCE SOME OF THE FIGURES DO NOT REPRODUCE WELL, INTERESTED PARTIES SHOULD CONSULT THE ORIGINAL SET. THANK YOU.

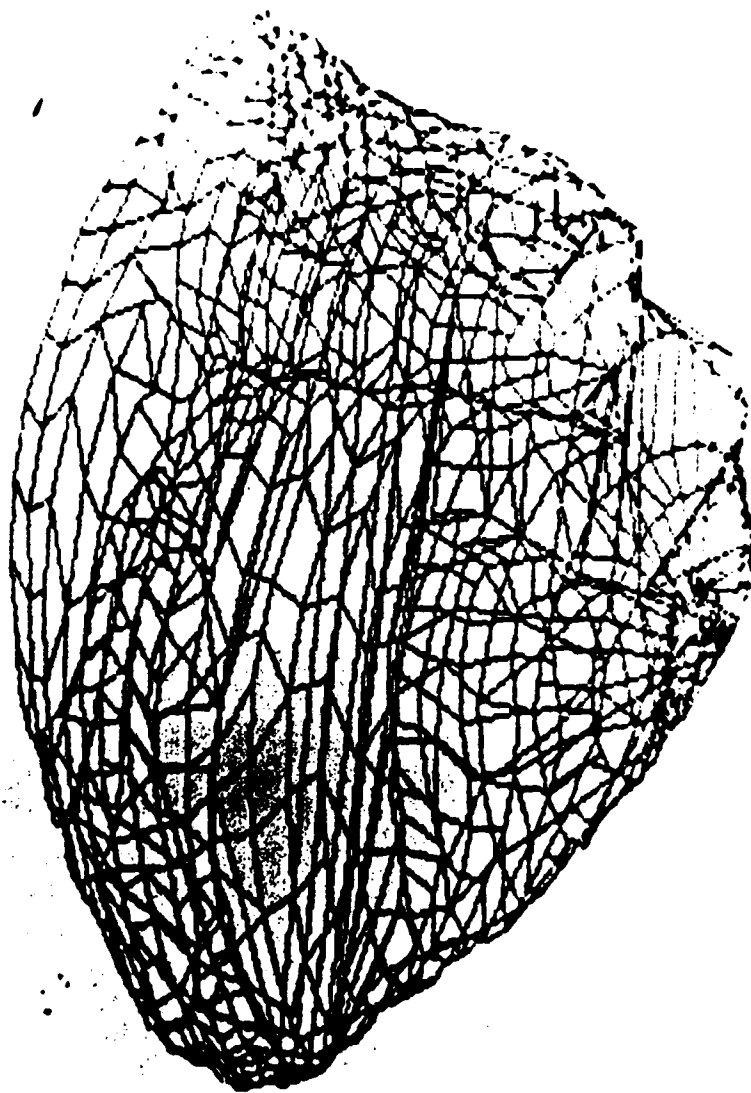
SWARTZ  
CHURCH  
AND  
ASSOCIATES

Swartz  
Figure 1





Schwartz  
Figure 2B



Schwartz  
Figure 3 A



Figure 3 B

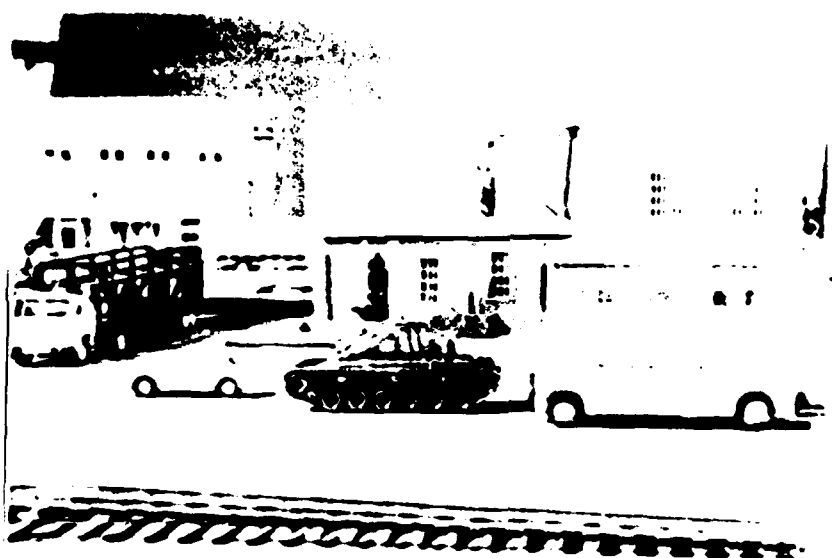
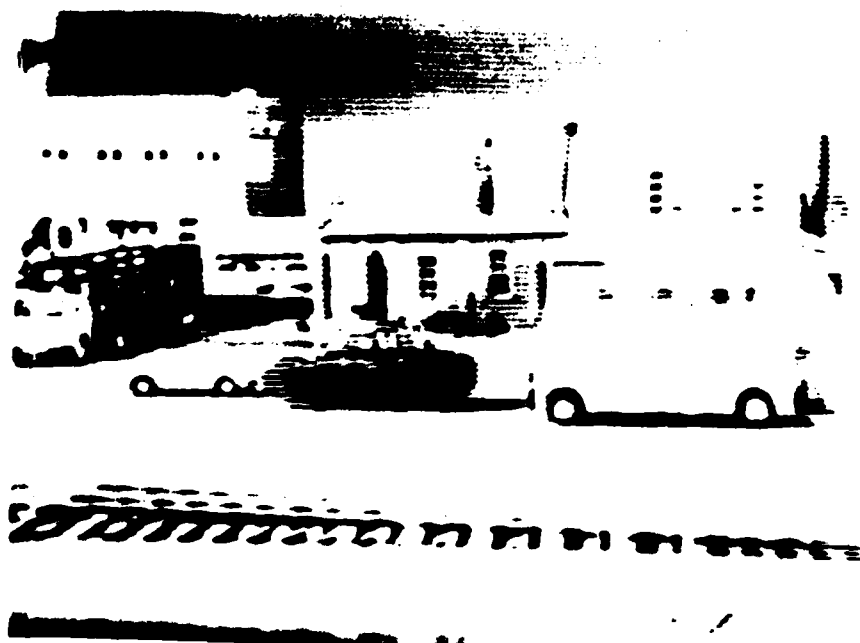
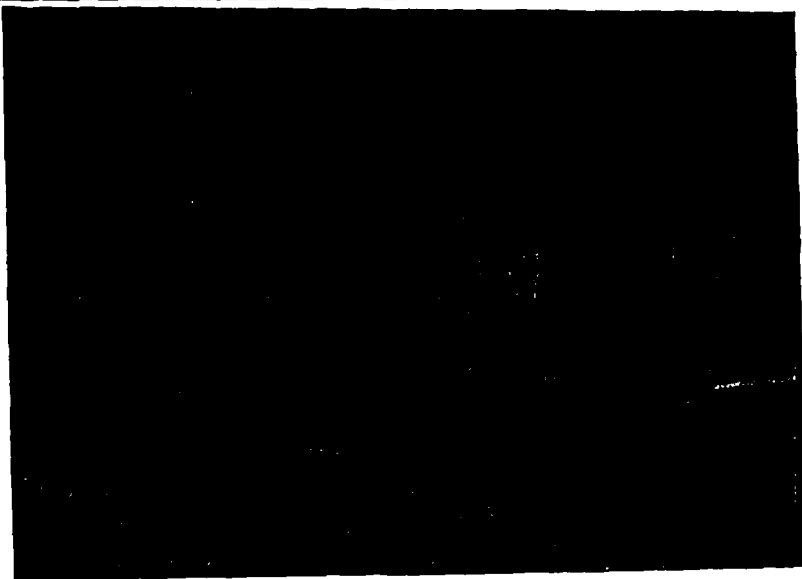


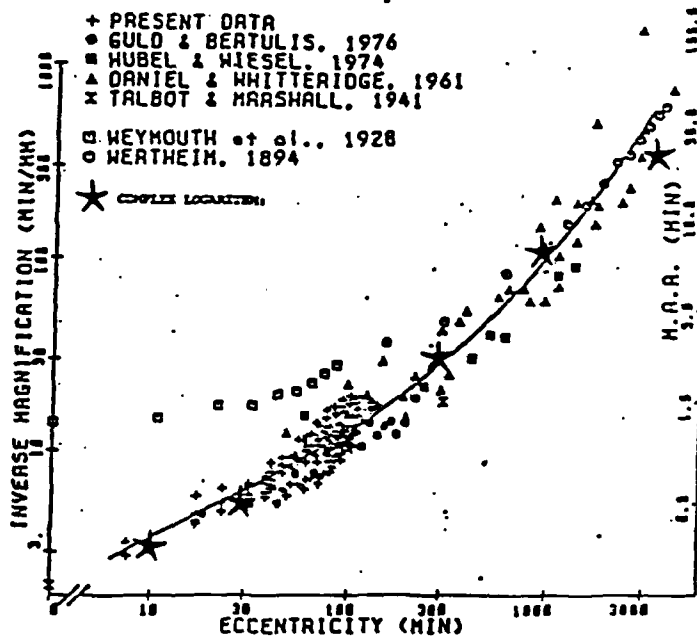
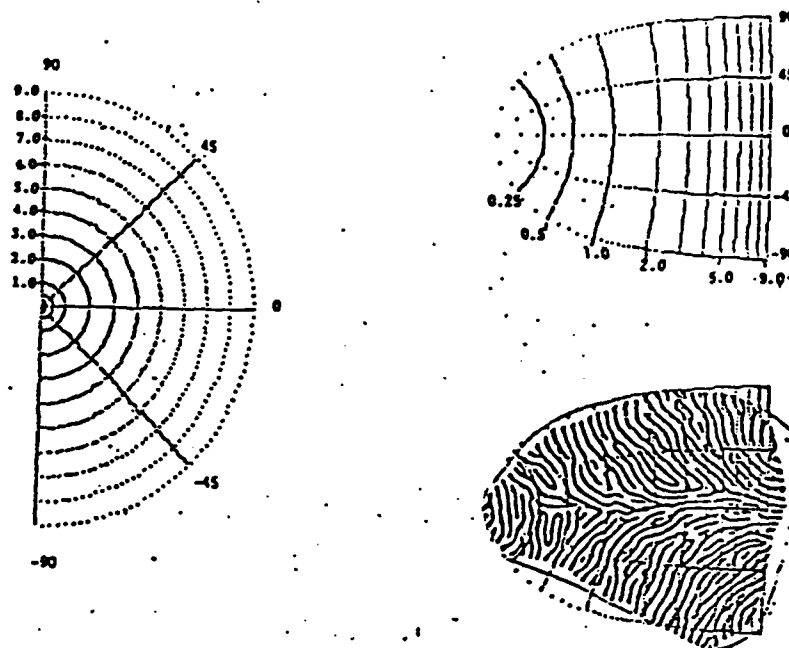
Figure 4





U N I C I T A T E

Schwartz  
Figure 6



Schwartz  
Figure 7





DATA COMPRESSION: LOG<sub>2</sub> 0.3  
MAP FUNCTION: LOG<sub>2</sub> 0.3  
FIELD OF VIEW: 8 DEGREE/HEMI-FIELD

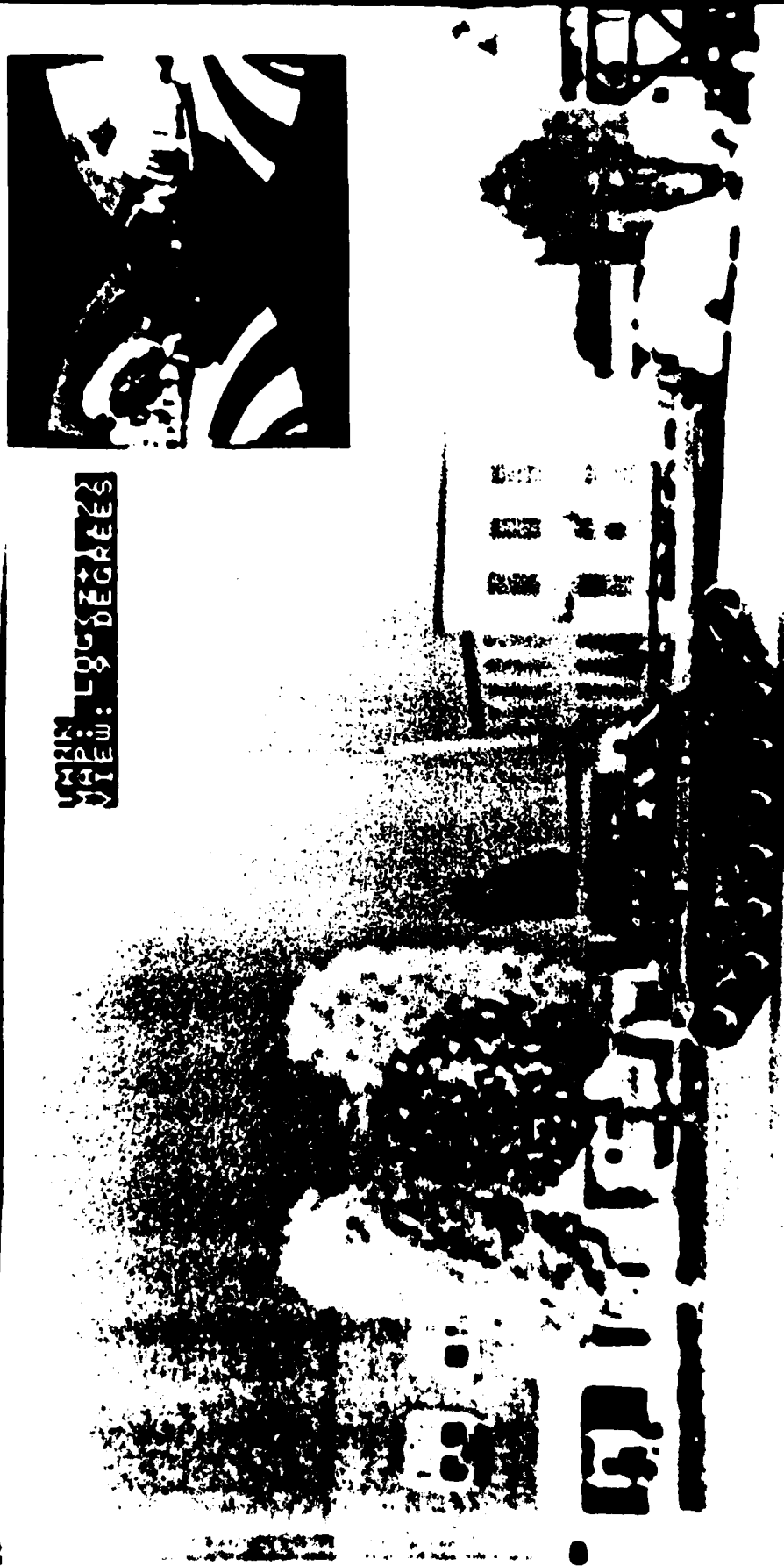


Schwartz  
Figure 8

Schwartz  
Figure 9

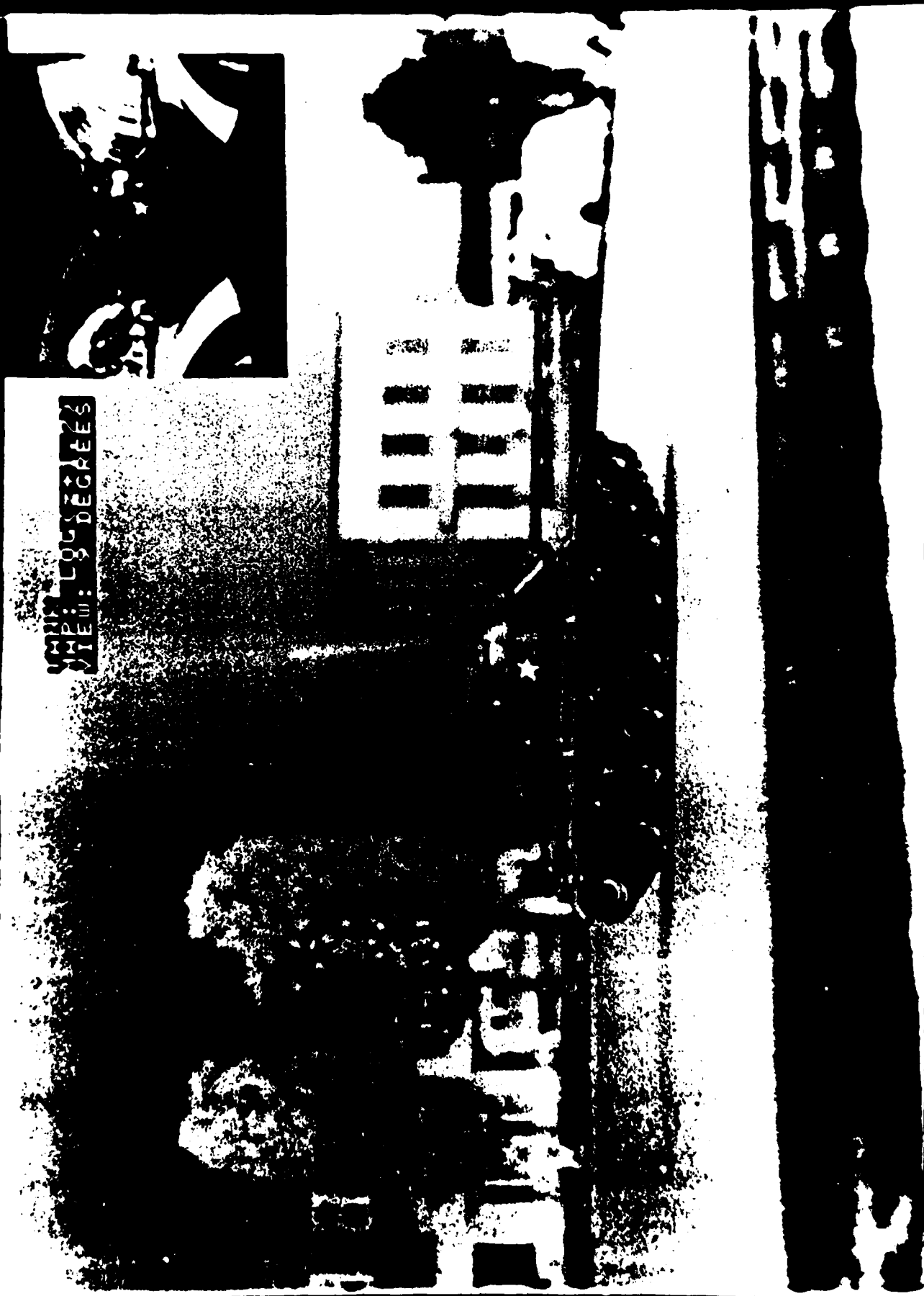


VIEW: LOG DEGREES

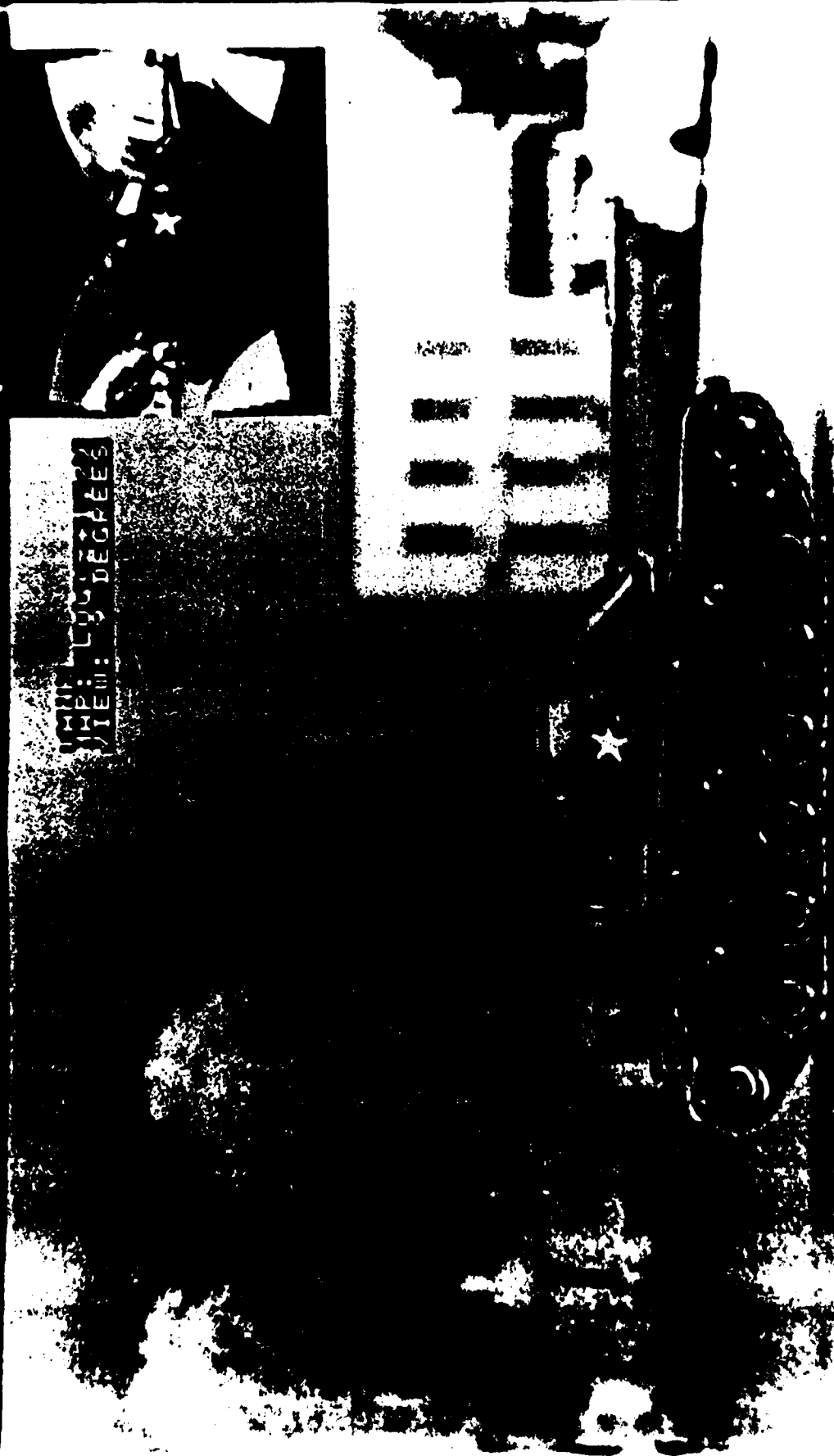


Schwartz  
Figure 10

THE  
NEW: 3 DEGREES

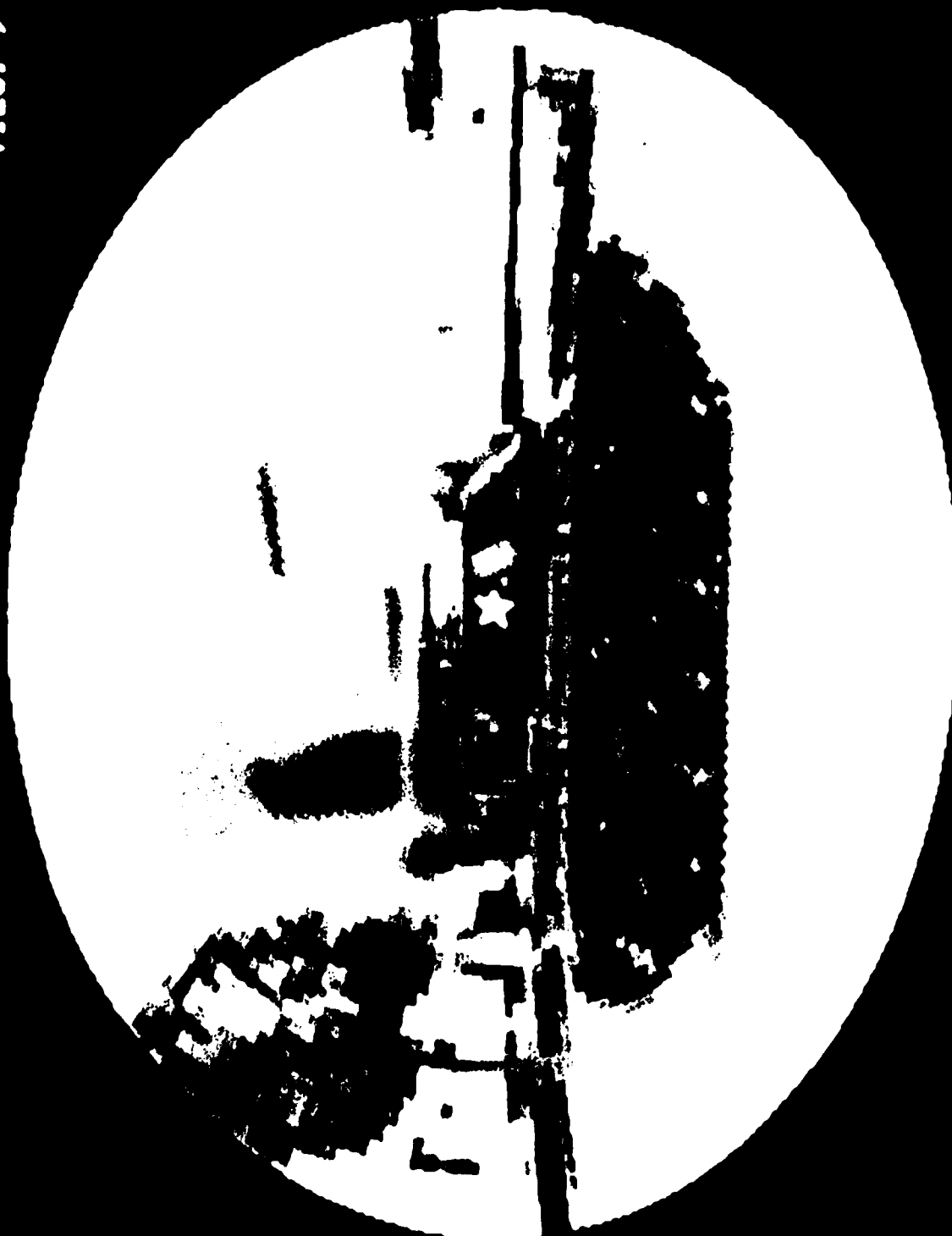


Schwartz  
Figure 11



Schwartz  
Figure 12

1991: 18 102-103



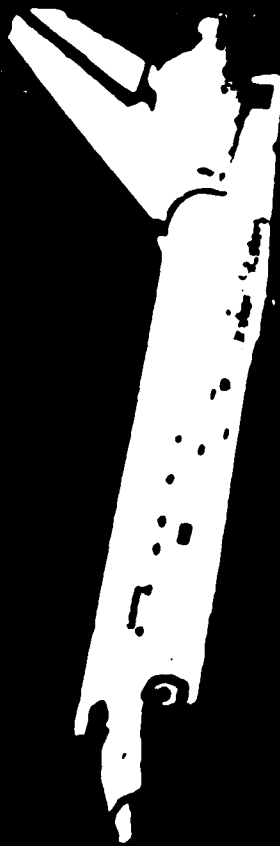
Schwartz  
Figure 13

TOPK 0910: 58' 24" 13

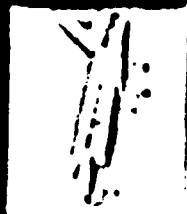
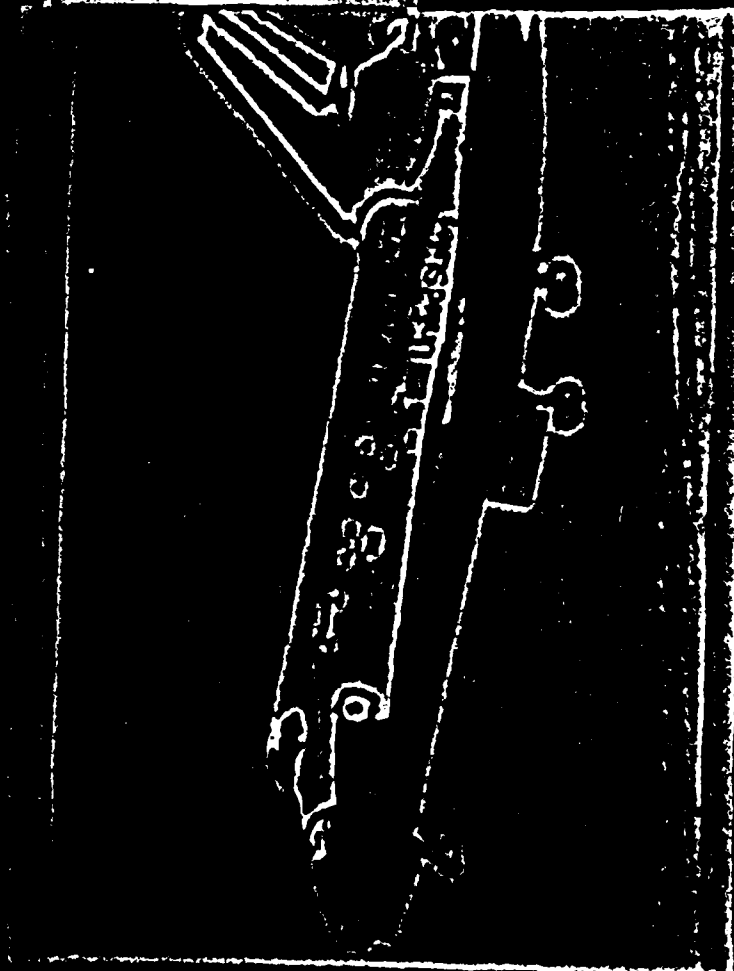
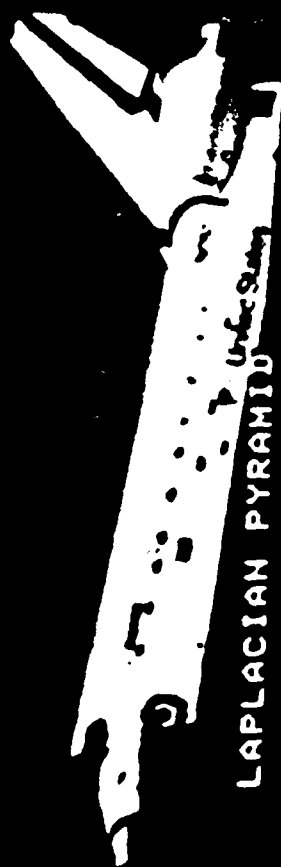


Schwartz  
Figure 14

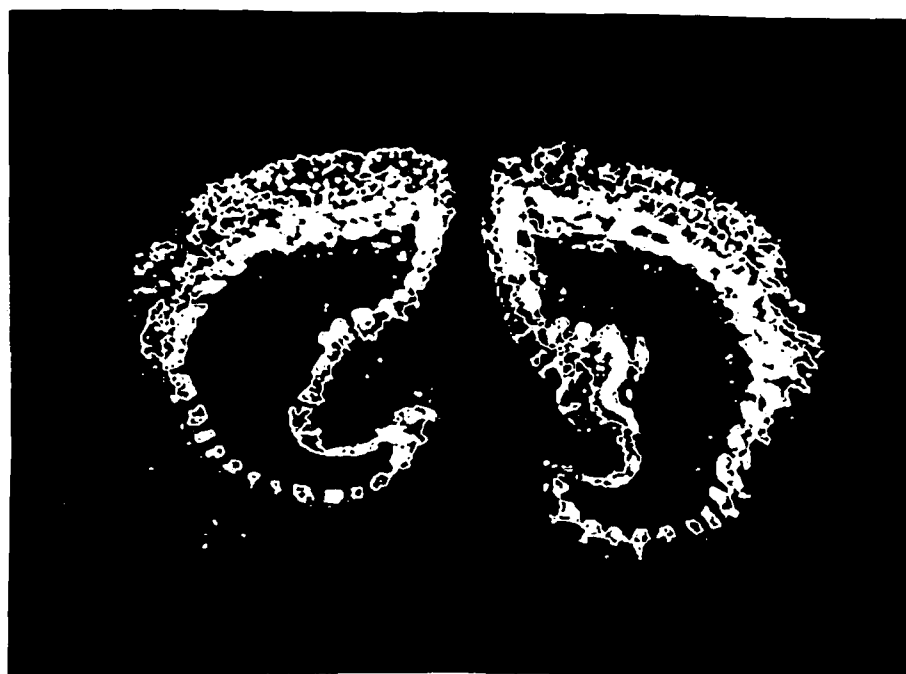
GAUSSIAN PYRAMID



Schwartz  
Figure 15







Schwartz  
Figure 17

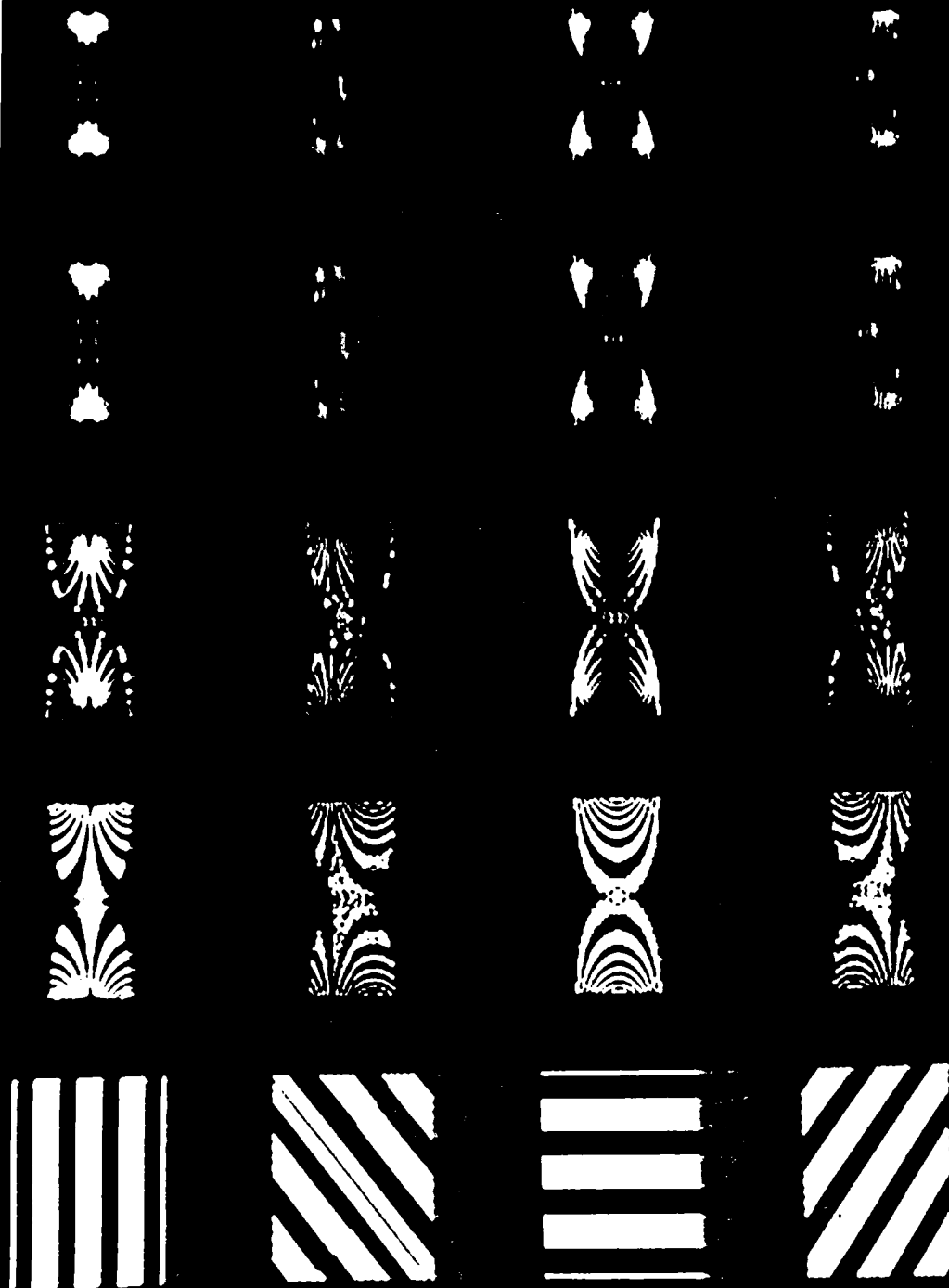


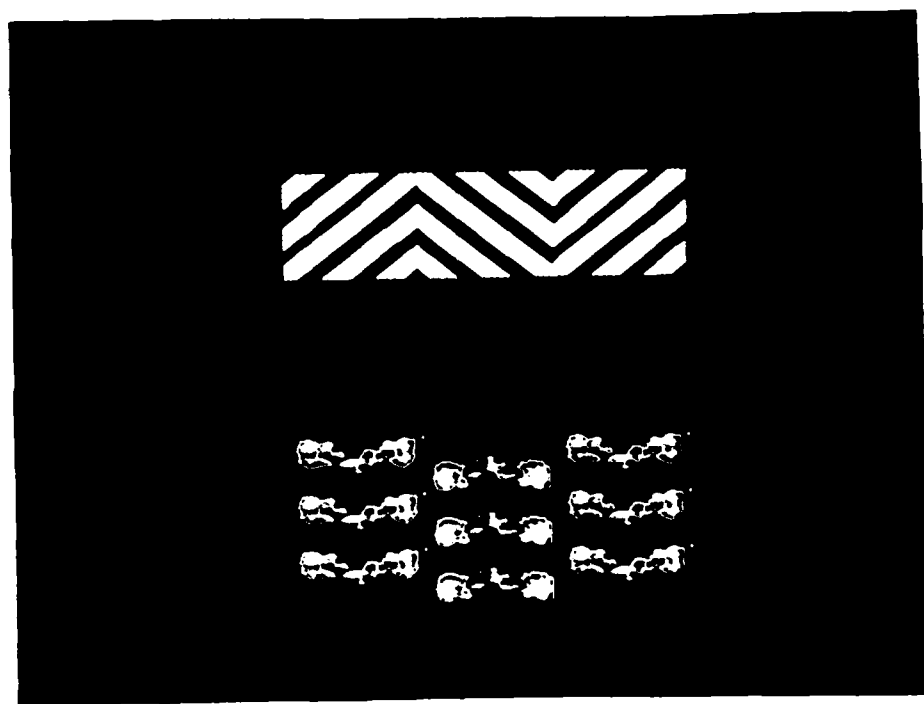
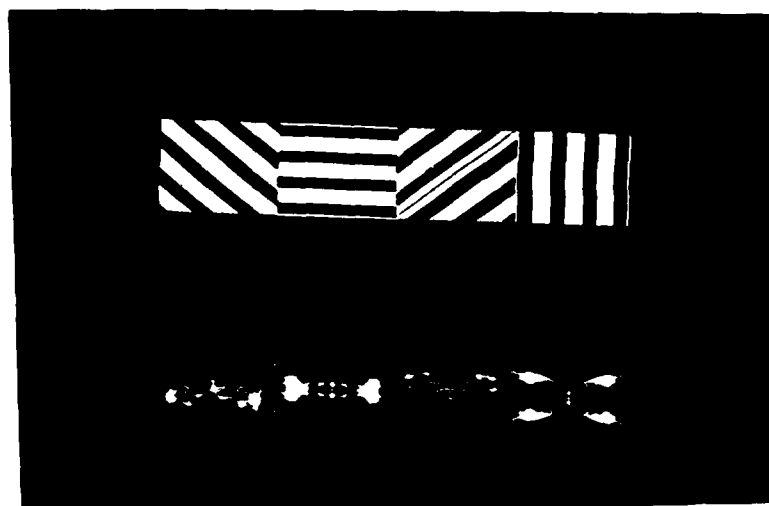
Schwartz  
Figure 18



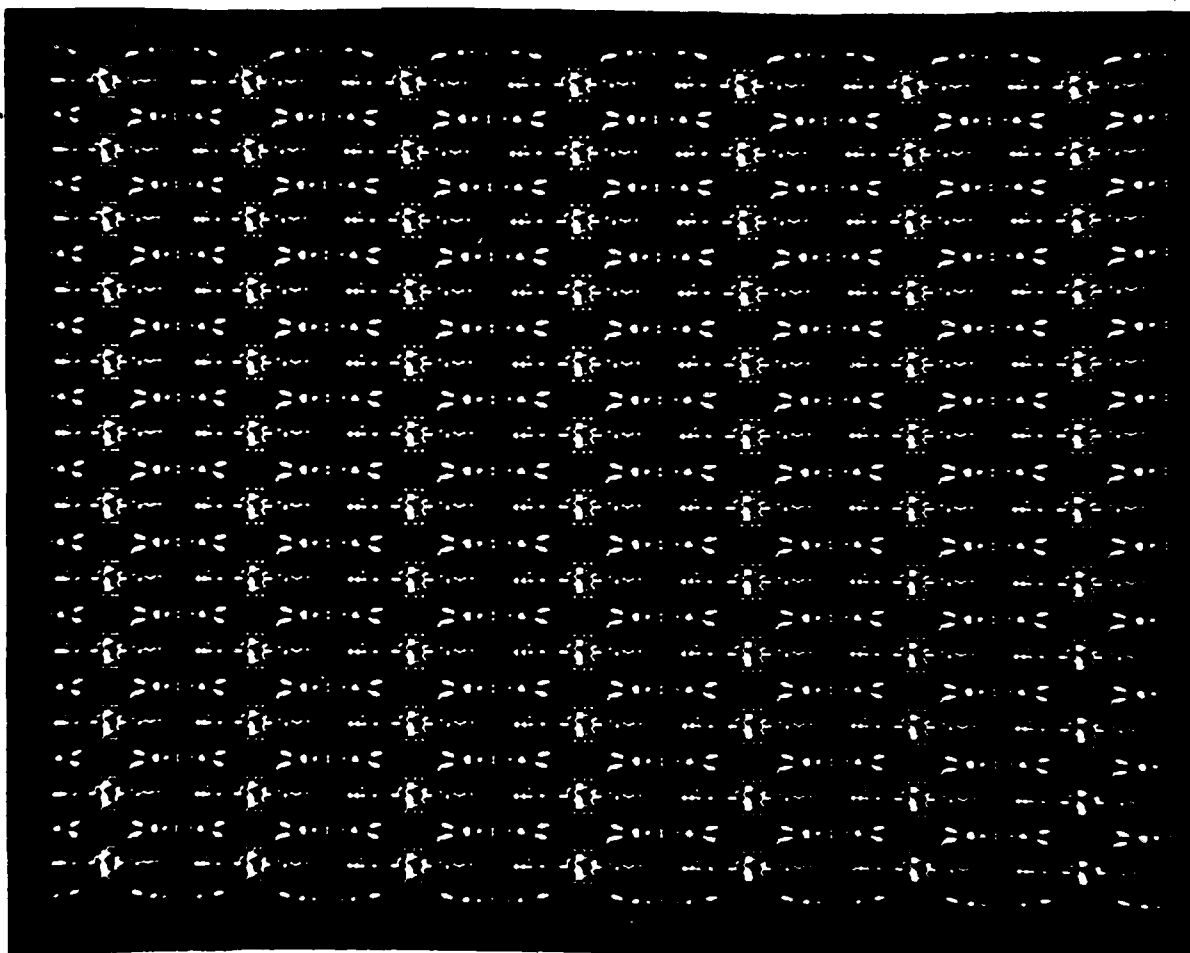


Schwartz  
Figure 20

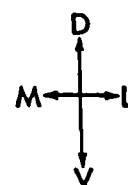
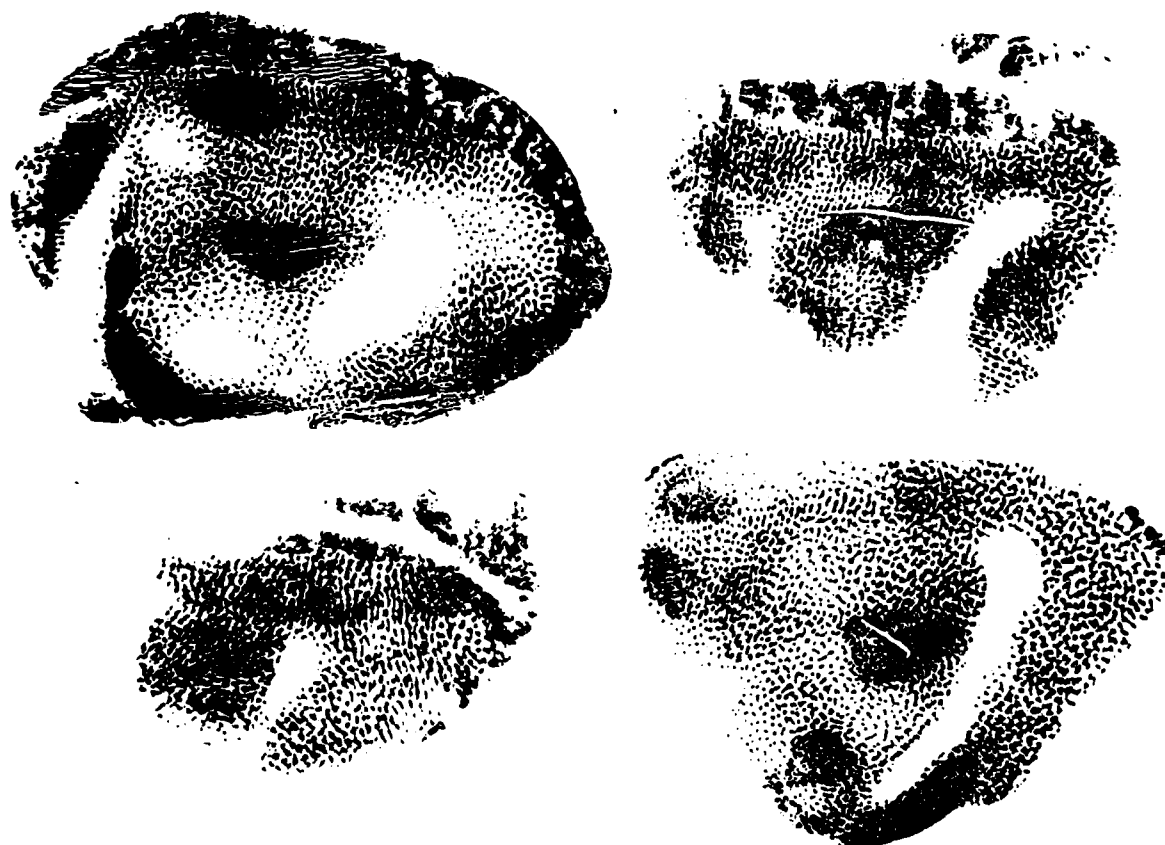




Schwartz  
Figure 22A



Schwartz  
Figure 22 B



1cm

Schwartz  
Figure 23

SPACE SHUTTLE LOG DEGREE 8.32  
FIELD OF VIEW: LOG DEGREE 8.32

MULTI-RESOLUTION PYRAMID ADD: 9 VIEW POINTS





Schwartz  
Figure 24



TOWN: LOG DEGREES  
VIEW: LOG DEGREES



TOWN: LOG DEGREES  
VIEW: LOG DEGREES



TOWN: LOG DEGREES  
VIEW: LOG DEGREES



TOWN: LOG DEGREES  
VIEW: LOG DEGREES

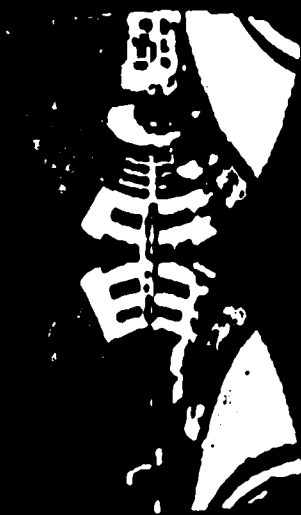


TOWN: LOG DEGREES  
VIEW: LOG DEGREES



TOWN: LOG DEGREES  
VIEW: LOG DEGREES

Schwartz  
Figure 25



TOWN: LOG 28+122  
VIEW: LOG 28+122



TOWN: LOG 28+122  
VIEW: LOG 28+122



TOWN: LOG 28+122  
VIEW: LOG 28+122



TOWN: LOG 28+122  
VIEW: LOG 28+122



TOWN: LOG 28+122  
VIEW: LOG 28+122



TOWN: LOG 28+122  
VIEW: LOG 28+122

Schwartz  
Figure 26



TOWN  
VIEW: LOGS+122



TOWN  
VIEW: LOGS+122



TOWN  
VIEW: LOGS+8.32



TANK  
VIEW: LOGS+122

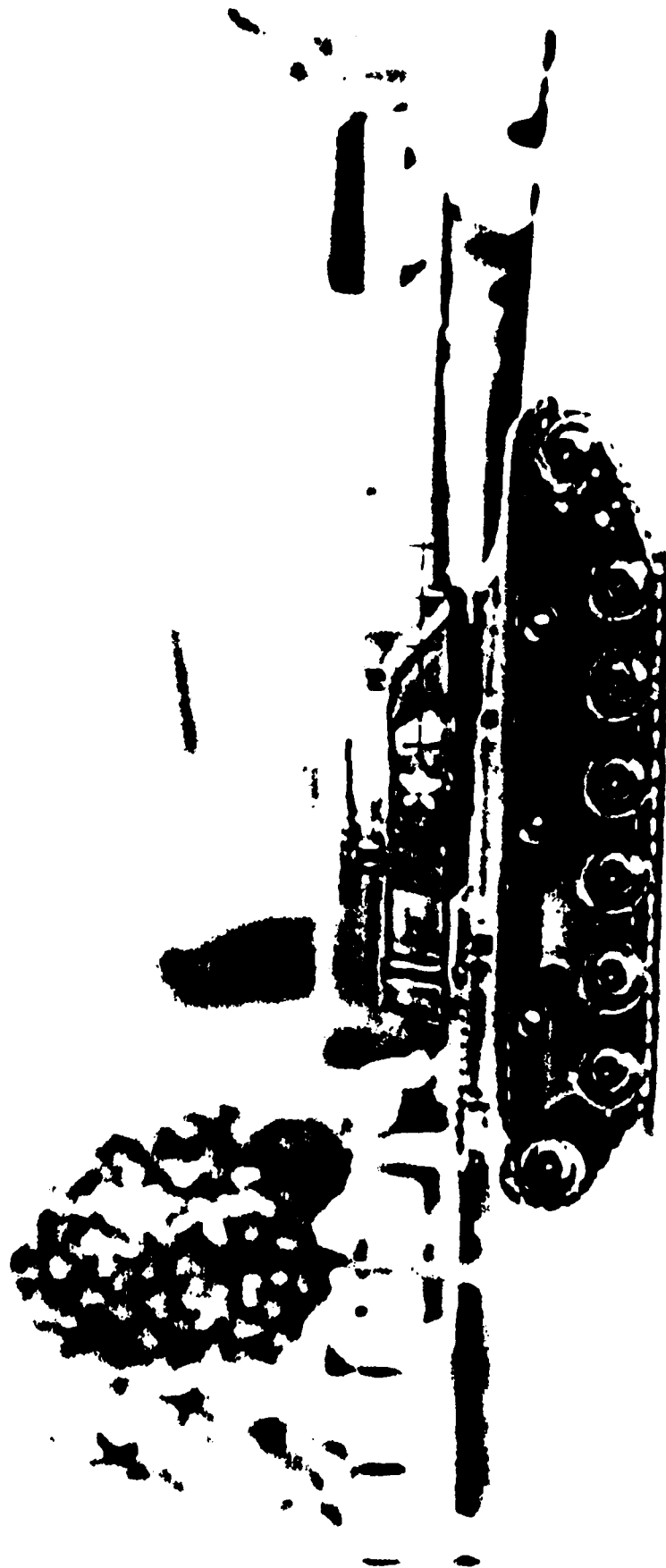


TANK  
VIEW: LOGS+122



TANK  
VIEW: LOGS+8.32

Schwartz  
Figure 27



Schwartz  
Figure 28



TANK OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK OFF ZOOM PT.  
VIEW: LOG DEGREES

Schwartz  
Figure 29



TANK: OFF ZOOM PT.  
VIEW: 0 DEGREES



TANK: OFF ZOOM PT.  
VIEW: 10 DEGREES



TANK: OFF ZOOM PT.  
VIEW: 20 DEGREES



TANK: OFF ZOOM PT.  
VIEW: 30 DEGREES



TANK: OFF ZOOM PT.  
VIEW: 40 DEGREES



TANK: OFF ZOOM PT.  
VIEW: 50 DEGREES

Schwartz  
Figure 30



TANK G. LOG OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK G. LOG OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK G. LOG OFF ZOOM PT.  
VIEW: LOG DEGREES



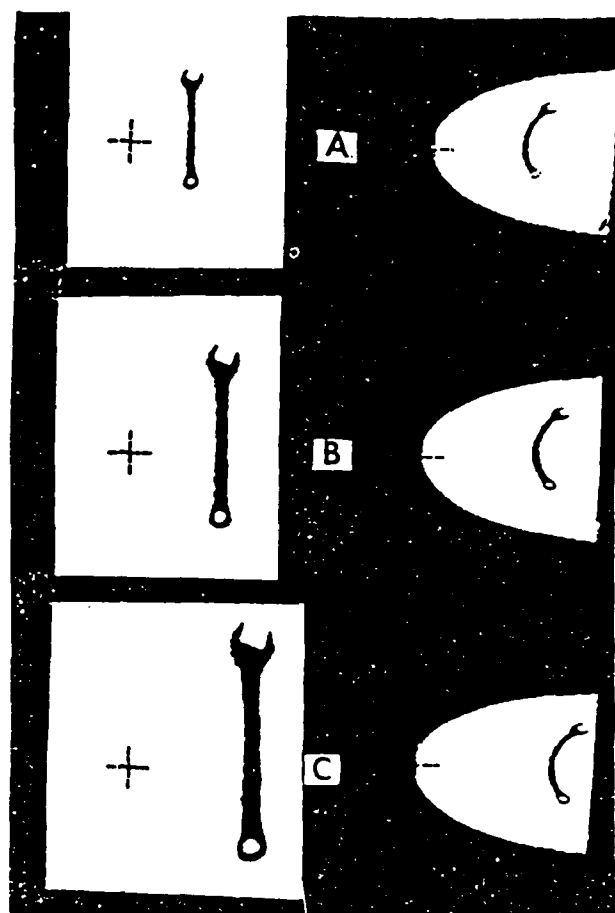
TANK G. LOG OFF ZOOM PT.  
VIEW: LOG DEGREES



TANK G. LOG OFF ZOOM PT.  
VIEW: LOG DEGREES

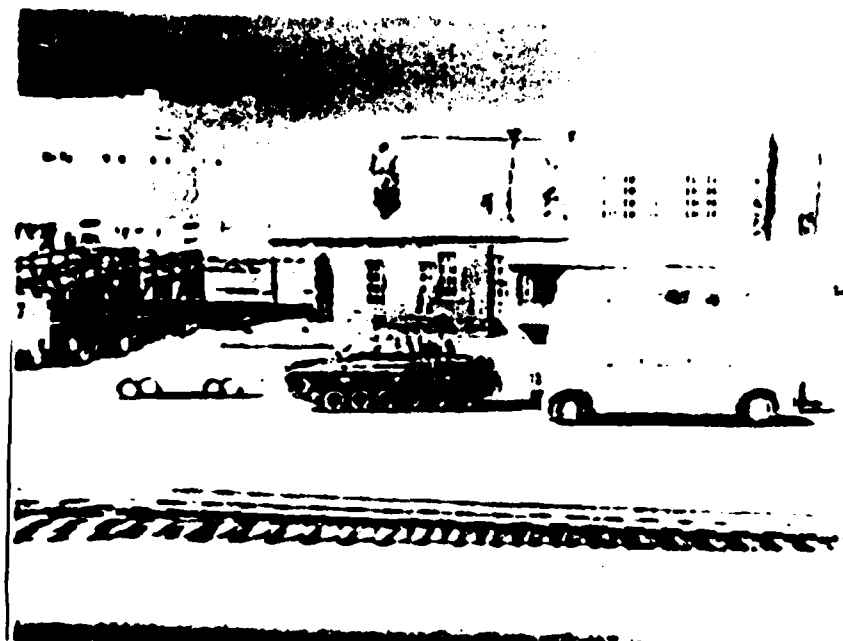


TANK G. LOG OFF ZOOM PT.  
VIEW: LOG DEGREES

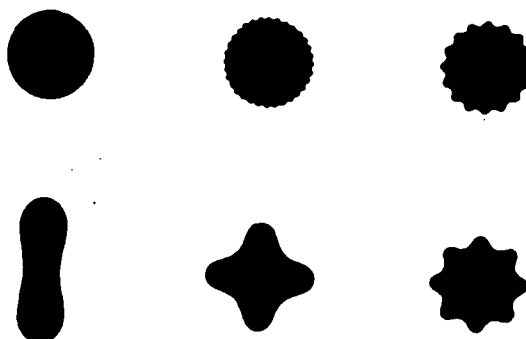


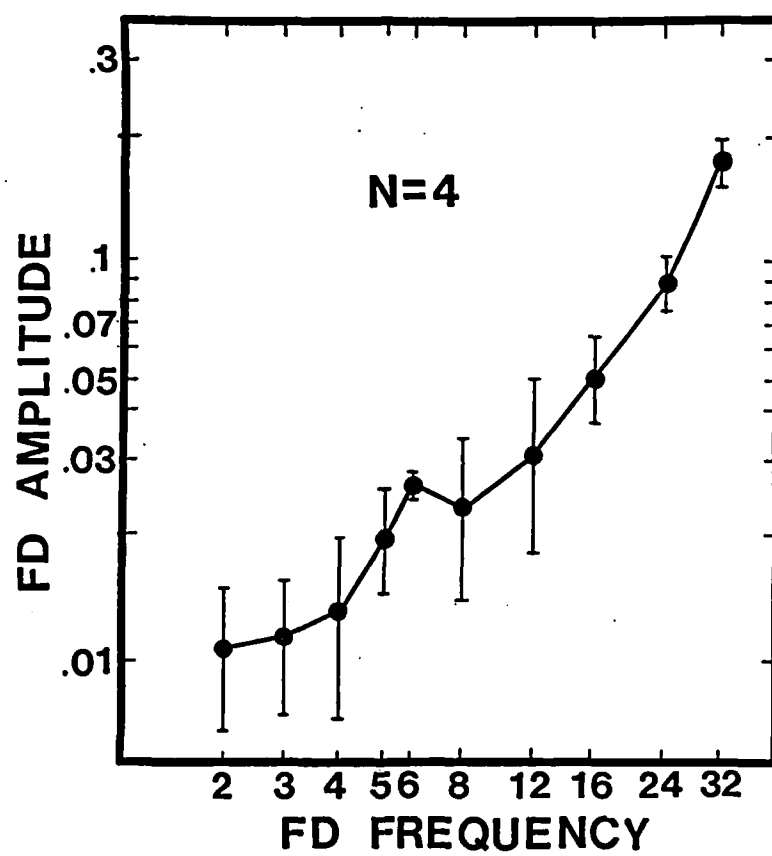


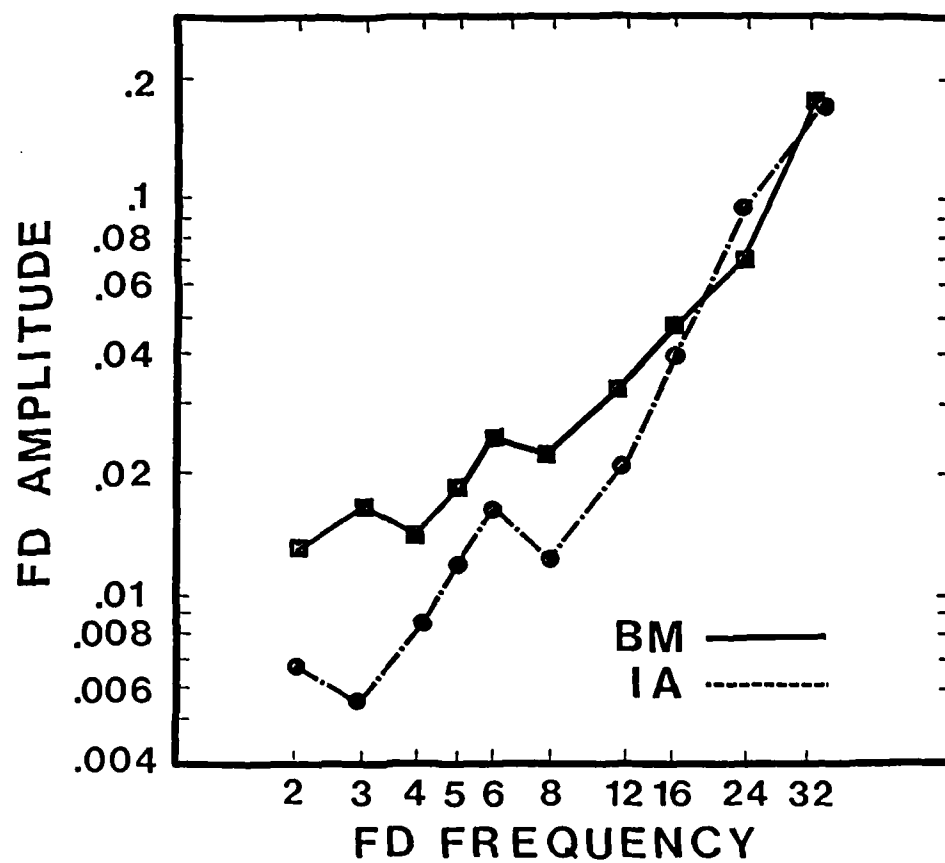
Schwartz  
Figure 32

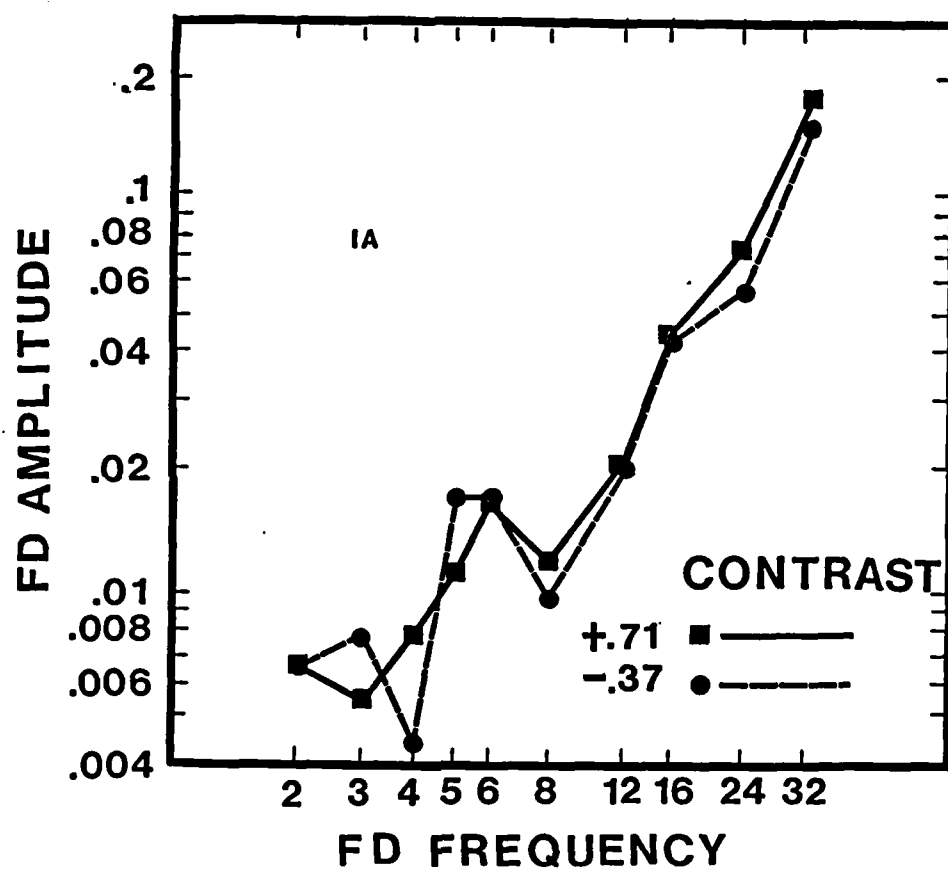


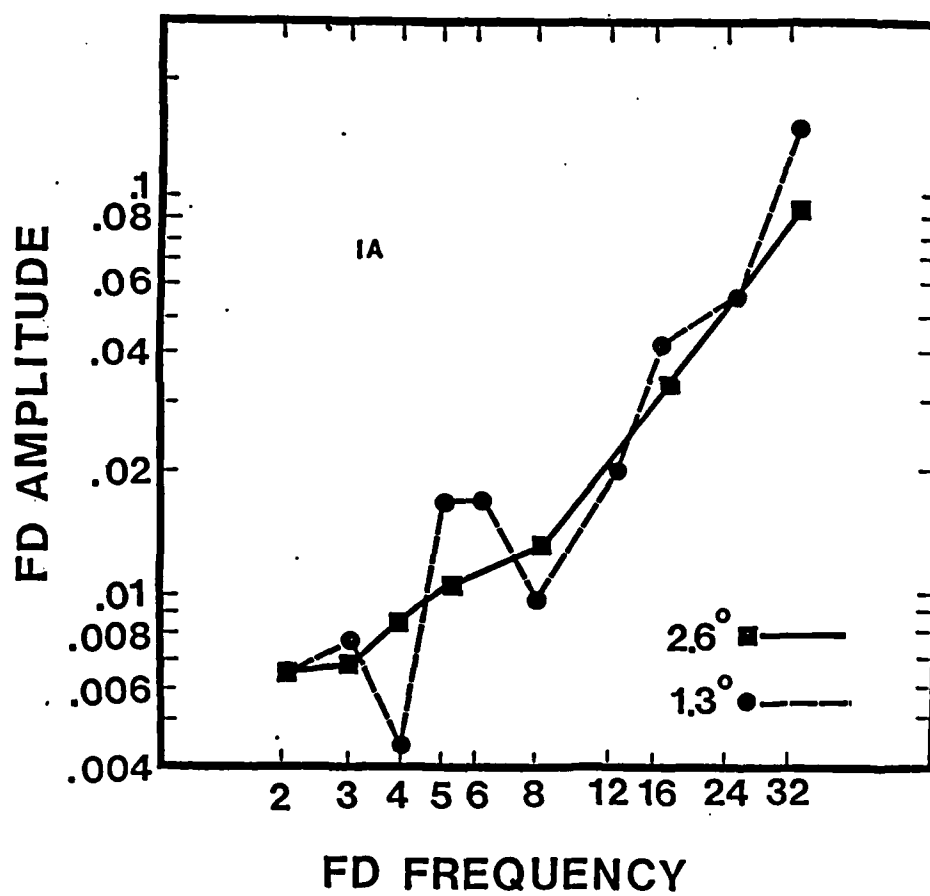
AMPLITUDE: 0.8

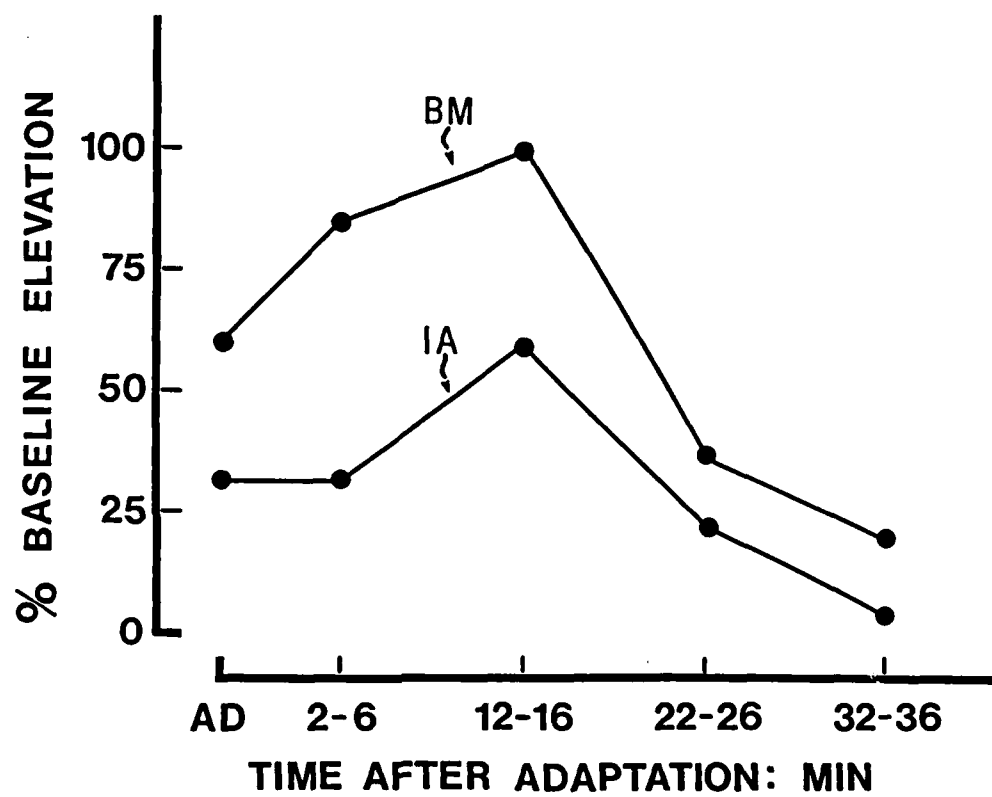


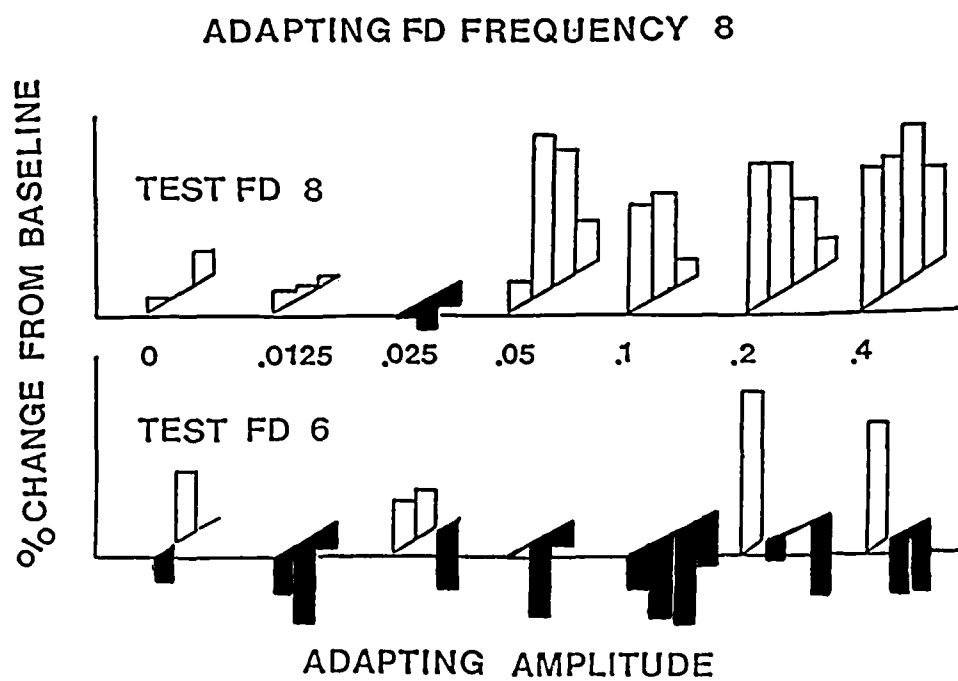














Schwartz  
Figure 40

